










Article

The Exploration of Predictors for Peruvian Teachers' Life Satisfaction through an Ensemble of Feature Selection Methods and Machine Learning

Luis Alberto Holgado-Apaza ^{1,*}, Nelly Jacqueline Ulloa-Gallardo ¹, Ruth Nataly Aragon-Navarrete ², Raidith Riva-Ruiz ³, Naomi Karina Odagawa-Aragon ⁴, Danger David Castellon-Apaza ¹, Edgar E. Carpio-Vargas ⁵, Fredy Heric Villasante-Saravia ⁵, Teresa P. Alvarez-Rozas ⁵ and Marleny Quispe-Layme ⁶

¹ Departamento Académico de Ingeniería de Sistemas e Informática, Facultad de Ingeniería, Universidad Nacional Amazónica de Madre de Dios, Puerto Maldonado 17001, Peru; nulloa@unamad.edu.pe (N.J.U.-G.); dcastellon@unamad.edu.pe (D.D.C.-A.)

² Departamento Académico de Ecoturismo, Universidad Nacional Amazónica de Madre de Dios, Puerto Maldonado 17001, Peru; raragon@unamad.edu.pe

³ Departamento Académico de Ciencias Económicas, Facultad de Ciencias Económicas, Universidad Nacional de San Martín, Tarapoto 22200, Peru; rriva@unsm.edu.pe

⁴ Escuela Profesional de Ingeniería de Sistemas e Informática, Facultad de Ingeniería, Universidad Nacional Amazónica de Madre de Dios, Puerto Maldonado 17001, Peru; nodagawa@unamad.edu.pe

⁵ Departamento Académico de Ingeniería Estadística e Informática, Universidad Nacional del Altiplano-Puno, Puno 21001, Peru; ecarpio@unap.edu.pe (E.E.C.-V.); fredy.villasante@unap.edu.pe (F.H.V.-S.); tpalvarez@unap.edu.pe (T.P.A.-R.)

⁶ Departamento Académico de Contabilidad y Administración, Facultad de Ecoturismo, Universidad Nacional Amazónica de Madre de Dios, Puerto Maldonado 17001, Peru; maquispe@unamad.edu.pe

* Correspondence: lholgado@unamad.edu.pe



Citation: Holgado-Apaza, L.A.; Ulloa-Gallardo, N.J.; Aragon-Navarrete, R.N.; Riva-Ruiz, R.; Odagawa-Aragon, N.K.; Castellon-Apaza, D.D.; Carpio-Vargas, E.E.; Villasante-Saravia, F.H.; Alvarez-Rozas, T.P.; Quispe-Layme, M. The Exploration of Predictors for Peruvian Teachers' Life Satisfaction through an Ensemble of Feature Selection Methods and Machine Learning. *Sustainability* **2024**, *16*, 7532. <https://doi.org/10.3390/su16177532>

Academic Editors: Pilar Colás-Bravo, Emilio Berrocal De Luna, Calixto Gutiérrez Braojos and Hao-Chiang Koong Lin

Received: 27 July 2024

Revised: 18 August 2024

Accepted: 28 August 2024

Published: 30 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Teacher life satisfaction is crucial for their well-being and the educational success of their students, both essential elements for sustainable development. This study identifies the most relevant predictors of life satisfaction among Peruvian teachers using machine learning. We analyzed data from the National Survey of Teachers of Public Basic Education Institutions (ENDO-2020) conducted by the Ministry of Education of Peru, using filtering methods (mutual information, analysis of variance, chi-square, and Spearman's correlation coefficient) along with embedded methods (Classification and Regression Trees—CART; Random Forest; Gradient Boosting; XGBoost; LightGBM; and CatBoost). Subsequently, we generated machine learning models with Random Forest; XGBoost; Gradient Boosting; Decision Trees—CART; CatBoost; LightGBM; Support Vector Machine; and Multilayer Perceptron. The results reveal that the main predictors of life satisfaction are satisfaction with health, employment in an educational institution, the living conditions that can be provided for their family, and conditions for performing their teaching duties, as well as age, the degree of confidence in the Ministry of Education and the Local Management Unit (UGEL), participation in continuous training programs, reflection on the outcomes of their teaching practice, work–life balance, and the number of hours dedicated to lesson preparation and administrative tasks. Among the algorithms used, LightGBM and Random Forest achieved the best results in terms of accuracy (0.68), precision (0.55), F1-Score (0.55), Cohen's kappa (0.42), and Jaccard Score (0.41) for LightGBM, and accuracy (0.67), precision (0.54), F1-Score (0.55), Cohen's kappa (0.41), and Jaccard Score (0.41). These results have important implications for educational management and public policy implementation. By identifying dissatisfied teachers, strategies can be developed to improve their well-being and, consequently, the quality of education, contributing to the sustainability of the educational system. Algorithms such as LightGBM and Random Forest can be valuable tools for educational management, enabling the identification of areas for improvement and optimizing decision-making.

Keywords: teacher satisfaction; teacher well-being; ensemble methods; machine learning; feature selection; prediction; educational sustainability; well-being

1. Introduction

Life satisfaction is an essential component in terms of achieving happiness and finding purpose in life [1]. This subjective concept reflects how a person perceives their own life [2,3]. Various international indicators offer perspectives on how life satisfaction varies across different countries. For example, the Better Life Index of the Organisation for Economic Co-operation and Development (OECD) compares well-being on a scale from 0 to 10. According to this index, the countries with the highest levels of life satisfaction are Finland (7.9), Iceland (7.6), Switzerland and the Netherlands (7.5), and Norway (7.3) [4]. On the other hand, the 2024 World Happiness Report highlights Finland (7.7), Denmark (7.6), Iceland (7.5), Sweden, and Israel (7.3) as the happiest countries, while Peru scores a relatively low value of 5.8 [5]. Another relevant indicator is the Human Development Index (HDI) of the United Nations Development Programme (UNDP), which classifies countries into development levels on a scale from 0 to 1. Countries with the highest HDI, such as Switzerland (0.97), Norway (0.97), Iceland (0.96), Hong Kong (0.96), and Sweden (0.95), tend to report higher levels of life satisfaction. In contrast, Peru, with an HDI of 0.76, is ranked at a high level but still below the highest indices [6]. These indicators highlight the importance of economic, social, and human development factors in life satisfaction.

Improving people's life satisfaction involves identifying the factors that influence it. Various studies have shown that demographic and socioeconomic factors, such as gender, marital status, and monthly income [1,7–10], play an important role. Additionally, psychological factors such as emotional intelligence, ego integrity, extroversion, conscientiousness, and job satisfaction [11,12] are also crucial. Participation in and enjoyment of leisure time are presented as aspects that positively influence life satisfaction [13]. Likewise, the balance between work and one's personal life [4,14], as well as physical factors such as a healthy lifestyle and both objective and subjective health status [15–17], significantly affects people's life satisfaction.

In the field of educational work, the challenges faced by teachers for various reasons affect the success of their important role as facilitators of learning, integral development, the promotion of critical thinking, and the development of social-emotional skills in students [18]. Teacher well-being, understood as the degree of satisfaction of teachers with the personal, social, and material conditions that allow them to perform their professional work in an optimal and meaningful way, is a crucial condition that needs to be identified and understood in order to design educational policies that promote healthy and productive work environments, which in turn improves educational outcomes and the well-being of society in general [11,18–20].

In recent years, machine learning (ML) has demonstrated its versatility and transformative potential across various fields, such as healthcare, engineering, occupational psychology, agriculture, finance, marketing, and transportation [21–46]. In the educational field, it has been used to evaluate students' academic performance and to identify and mitigate academic risk [47–50]. However, despite the growing interest in studying life satisfaction using ML, there is a notable lack of studies specifically focusing on the life satisfaction of basic education teachers. To date, no study has been found that investigates the life satisfaction of these teachers while simultaneously considering a wide range of predictors.

In light of this, the present study aims to identify the most significant predictors of life satisfaction among Peruvian teachers. To achieve this, we analyzed data provided by the National Survey of Public Basic Education Teachers (ENDO-2020), supplied by the Ministry of Education of Peru [51], using an ensemble of feature selection methods that combines both filtering and embedded techniques. Additionally, we assessed the suitability

of the selected features for predicting teachers' life satisfaction across eight cutting-edge ML algorithms. Finally, we identified the most optimal algorithms for predicting the life satisfaction of Peruvian teachers. This approach will not only enhance the understanding of factors affecting teachers' well-being, but also provide valuable tools for designing interventions and educational policies that promote a healthier and more satisfying work environment for educators.

The remainder of the article is organized as follows: Section 2 reviews the relevant literature. Section 3 details the methodology employed in the experiments. Section 4 presents the results obtained. Section 5 discusses these results, and finally Section 6 provides our conclusions.

2. Literature Review

2.1. Concept of Life Satisfaction and Influencing Factors

Life satisfaction is a widely explored topic in psychology and refers to individuals' overall evaluation of their life as a whole [52]. It is an evaluative process in which people judge the quality of their life based on a personal set of criteria [53,54].

Numerous studies have demonstrated that economic income is positively correlated with life satisfaction [9,55,56]. The balance between work and personal life also emerges as a crucial variable [57,58]. Additionally, a healthy lifestyle and a positive attitude towards well-being have been found to significantly impact life satisfaction [59].

In the context of teachers' life satisfaction, several factors play a significant role. Job satisfaction [60,61] and workload [62], for example, significantly influence teachers' overall satisfaction. A study conducted in Switzerland revealed that dispositional optimism, burnout, and job satisfaction are key determinants of teachers' life satisfaction [63]. Additionally, research in Italy found that progress towards personal work goals and positive affectivity are important predictors of teachers' life satisfaction [61]. Occupational stress has also been observed to be associated with lower life satisfaction among teachers in Pakistan [64]. A study in Peru demonstrated that school leadership positively influences teachers' life satisfaction [65]. Finally, research in Spain revealed a strong relationship between perceived emotional intelligence and teachers' life satisfaction.

2.2. Machine Learning Techniques and Their Application in the Study of Life Satisfaction

Machine learning (ML) is a subfield of artificial intelligence (AI) that focuses on the development of algorithms that enable computers to automatically identify patterns in data and improve their performance through experience, without explicit programming [66]. AI, on the other hand, is the broadest field within computer science. It encompasses a range of technologies and methods designed to enable machines to think and act in ways similar to humans.

The literature shows that various approaches have been explored when using ML techniques to analyze life satisfaction, with the focus primarily on predictive modeling and the analysis of influencing factors. ML has proven to be an effective tool for this purpose. Byeon [67] developed a predictive model for elderly individuals in South Korea using the classification and regression trees (CART) algorithm, identifying significant variables such as subjective perceptions of friendship, health status, family relationships, and educational level. Zhang and Li [68] applied Support Vector Machines to predict life satisfaction in elderly individuals, combining factor analysis with a Particle Swarm Optimization algorithm to achieve an accuracy of 94.23%. Pan and Cutumisi [69] used Random Forest and K-Nearest Neighbors with data from secondary school students in the United Kingdom and Japan, based on the 2018 Programme for International Student Assessment (PISA). They found that Random Forest outperformed K-Nearest Neighbors and highlighted factors such as the meaning of life, student competence, teacher support, exposure to bullying, and ICT resources as important in predicting students' life satisfaction. Khan et al. [70] demonstrated the effectiveness of Random Forest in predicting life satisfaction in Denmark, achieving an accuracy of 93.80% and a macro F1 score of 70.60%, and emphasized the

relevance of explainable artificial intelligence (XAI) in researching subjective well-being. Finally, Jaiswal et al. [71] built a model to predict overall happiness and found that Random Forest provided the highest accuracy at 92.27%, revealing a notable difference in satisfaction between countries based on GDP per capita.

Another relevant approach in studying life satisfaction is using ML techniques to analyze influencing factors. Morrone et al. [72] employed ordinal classification trees, accounting for both ordinal and diverse satisfaction categories, and found that greater disadvantages are linked to lower satisfaction levels. Lee [73] used Random Forest methods to analyze the 2021 Youth Socioeconomic Survey's data from South Korea, identifying 21 key variables, including work values, that affect youth life satisfaction. Shen et al. [74] explored various ML algorithms to predict life satisfaction in older adults, highlighting that subjective social status, positive emotions, and negative emotions are critical predictors. Kim et al. [16] examined how leisure-related factors affect happiness in individuals over 65 in South Korea, finding that satisfaction with leisure, free time, and public facilities are significant predictors, with Support Vector Machine being the most effective algorithm. Finally, Jang and Masatsuku [75] analyzed the happiness of Korean residents through topic modeling and neural networks, revealing that variables such as family life, income, social status, health, and perceived inequality are keys to happiness.

2.3. Ensemble of Feature Selection Methods

These approaches are based on the principle that integrating the results of multiple feature selection algorithms can be more effective than relying on a single algorithm [76,77]. The ensemble of feature selection methods utilizes the concept of model ensembling in machine learning, where multiple feature selectors are employed in an initial phase and their results are combined in a later phase to produce a unified final outcome [78]. Several studies have shown that ensembles of feature selection methods often improve performance metrics in classification problems compared to the individual use of a single selection method [79–84].

3. Materials and Methods

The machine learning approach used in exploring predictors of teachers' life satisfaction focuses on identifying relevant features from the database of the National Survey of Teachers of Public Educational Institutions of Regular Basic Education (ENDO) and validating them using metrics obtained during the testing phase of the models. Figure 1 illustrates the flow chart of our methodological approach, which consists of four steps: (1) data extraction, (2) data cleaning and preprocessing, (3) feature selection, and (4) training and model evaluation.

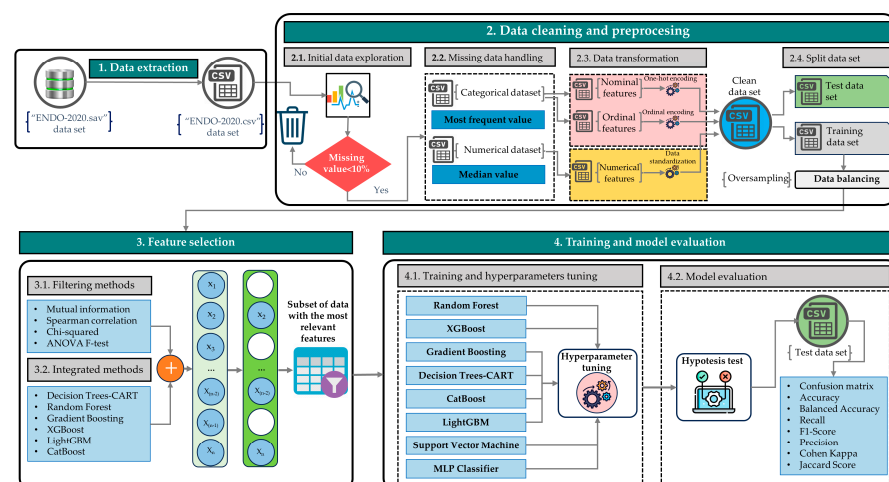


Figure 1. Methodological proposal for exploration of predictors of life satisfaction among elementary school teachers in Peru.

3.1. Data Extraction

The data for the present research were obtained from ENDO, available on the official website of the Peruvian Ministry of Education [51] ENDO-2020 was applied remotely (via telephone) from 24 November to 1 December 2020 to a sample of 28,127 teachers from regular basic education institutions in all regions of Peru. This survey collects updated information from teachers throughout the country, covering socio-demographic and socioeconomic characteristics, training, and professional careers. It also includes teacher perceptions about the working conditions that affect their well-being, their perceptions of policies and programs promoted by the MINEDU that impact their work, as well as their values and expectations for the future.

3.2. Data Cleaning and Preprocessing

Data cleaning and preprocessing are fundamental tasks in the work of the data scientist, with significant implications for the performance of ML techniques [85,86]. Currently, it is estimated that these tasks take up 60% to 80% of the total development time of a data science project [87,88]. This proportion varies depending on the initial quality of the data, as well as their complexity and nature. In the following, we detail the tasks performed in this phase.

3.2.1. Initial Data Exploration

We performed this task to obtain an overview of the dataset's characteristics, including its size, the amount of missing data, the presence of outliers, and the distribution of variables by data type. As observed in Table 1, the dataset consists of 150 variables and 28,216 rows. A total of 1.5265×10^6 missing cells were found, representing 36.1% of the total. No duplicate rows were found. The variables are distributed into 125 categorical and 25 numerical variables.

Table 1. General summary of the “ENDO-2020” dataset.

Attribute	Value Obtained
Variables	150
Rows	28,216
All missing cells	1.5265×10^6
Missing cell (%)	36.1%
Duplicate rows	0
Duplicate rows (%)	0.0%
Variable types	Categorical: 125 Numerical: 25

3.2.2. Missing Data Handling

In this task, we meticulously selected 92 columns, excluding those that contained a single value in all rows and those with more than 10% missing values. These exclusions were made to avoid potential errors or unforeseen results in the analysis, as suggested in previous studies on the impact of incomplete data on analysis quality [89,90].

Missing values were imputed using the mode for categorical variables and the median for numerical variables, methods commonly accepted for their computational simplicity [91,92]. Figure 2a shows an overview of the variables before imputation, while Figure 2b shows the variables after the imputation process.

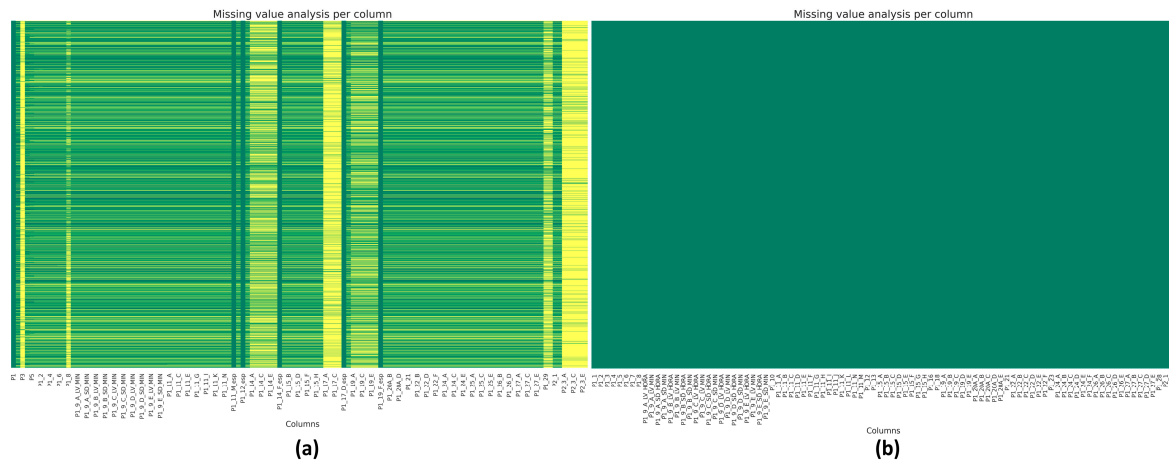


Figure 2. Report on missing data treatment: (a) dataset before imputation process; (b) dataset after imputation process. Yellow represents missing data. Green represents existing data.

3.2.3. Data Transformation

Given that the dataset consisted of 38 categorical columns, of which 15 were nominal and 23 were ordinal, in addition to 54 numerical columns, it was necessary to apply various encoding and normalization techniques, as most ML algorithms do not support categorical data [93,94]. To represent the 15 nominal categorical variables, the One Hot Encoding technique was employed, converting each category into a new binary column indicating the presence or absence of that category in each observation [16,95]. On the other hand, the 23 ordinal categorical variables were transformed using Ordinal Encoding, assigning each category an ordered numerical value that preserved the natural order of the categories [96]. Additionally, the numerical columns were standardized to ensure that the values had a distribution with a mean of 0 and a variance of 1 [97]. The implementations of these techniques were carried out using the OneHotEncoder, OrdinalEncoder, and StandardScaler classes from the preprocessing module of the Scikit-learn library in Python [98]. After this task, we obtained a dataset with 18,911 rows and 121 columns.

3.2.4. Split Dataset

We divided the dataset obtained in the previous phase into an 80:20 ratio, that is, we used 80% for the model training process and 20% for the testing process. This practice is common in studies of this nature, as observed in previous works [70,74]. The division was performed completely randomly to avoid potential biases that could arise from the order of the data. In this way, we obtained 15,128 instances in the training dataset and 3783 in the testing dataset.

Due to the imbalance in the target variable “teacher’s life satisfaction”, we balanced the data in the training dataset using the oversampling technique [99]. This technique involves creating additional copies of the minority class instances until a balance with the majority class is achieved. For this, we used the RandomOverSampler class [100], which is part of the imblearn.over_sampling module in Python. Figure 3 shows the class distribution of the target variable in three stages: (a) the training dataset, where an initial imbalance is observed; (b) the test dataset, which reflects the same imbalance; and (c) the test dataset after data balancing, where the oversampling technique has balanced the class representation. When a classifier works with an imbalanced dataset, it tends to show bias, achieving a higher accuracy for the majority class but an inferior performance for the minority class [101]. In this context, balancing is crucial to improve the performance of classification models and ensure a more equitable evaluation [102].

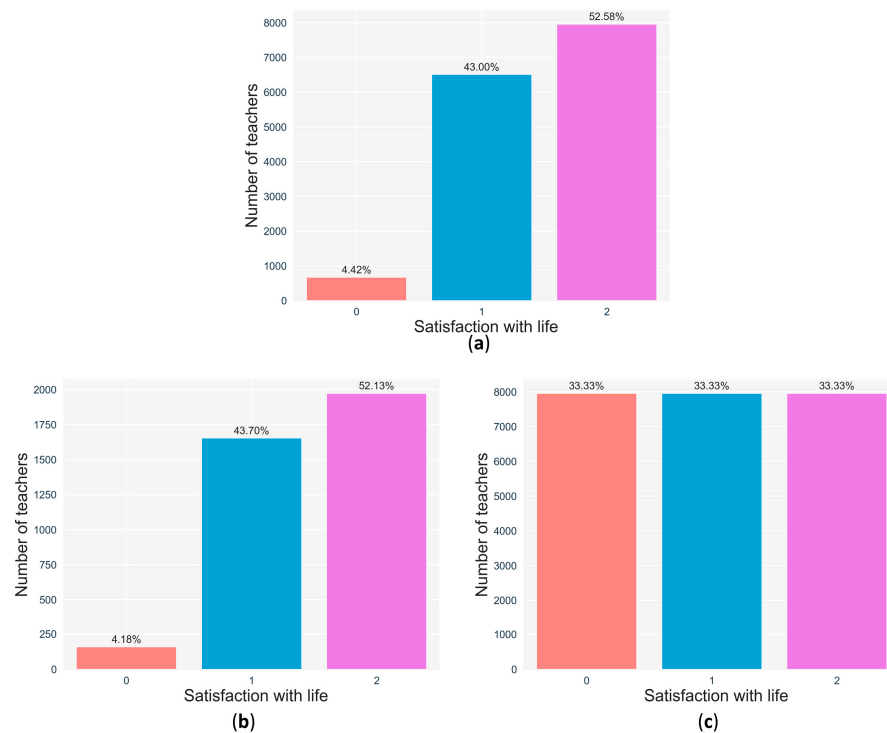


Figure 3. Distribution of the classes in the target variable: (a) training dataset; (b) test dataset; (c) training dataset after data balancing.

Since the target variable “teacher life satisfaction” was imbalanced, we performed data balancing in the training dataset using the oversampling technique [103]. This technique involves creating additional copies of instances from the minority class until a balance is achieved with the majority class. For this purpose, we utilized the `RandomOverSampler` class [100] from the `imblearn.over_sampling` module in Python. Figure 3 illustrates the class distribution for the target variable at three stages: (a) the training dataset, where an initial imbalance is observed; (b) the testing dataset, which reflects the same imbalance; and (c) the training dataset after data balancing, where the oversampling technique balances the class representation. When a classifier works with an imbalanced dataset, it tends to show bias, achieving higher accuracy in the majority class but a lower performance in the minority class [101].

3.3. Feature Selection

Feature selection in ML is a key process that identifies an optimal subset of relevant features, thereby improving model performance, reducing noise, and computational complexity [104]. By eliminating redundant or irrelevant features, the model’s interpretability is increased and more efficient and robust models are constructed [105]. The three most commonly used feature selection methods are filter-based methods, wrapper-based methods, and embedded methods [106–108]. In this research, we combined filter-based methods and embedded methods, using the stratified cross-validation technique with 10 folds.

3.3.1. Feature Selection by Filtering Methods

These techniques evaluate the relative value of each feature with respect to the target variable on their own, independently of any specific learning model [106,109–111]. To determine the quality of the feature subset or rate each variable according to a relevance index, researchers use statistical techniques, excluding those that fall below a predetermined threshold [76]. The analysis of variance (ANOVA), mutual information, correlation coefficient, and chi-square test are some of the most popular filtering techniques [110]. In this

research, we used the following filter methods: mutual information, ANOVA, chi-square, and Spearman's Correlation.

- Mutual information (*MI*) is a metric that quantifies the dependence between two variables, indicating to what extent knowledge of a feature helps predict the target variable [112]. Its value varies between 0 and 1, where 0 indicates no dependence and a value of $MI > 0$ indicates some relationship between the feature and the target variable [113]. For the selection of the k best features using the mutual information filter, we set the parameters `score_func = mutual_info_classif` and `k = 'all'` in the `SelectKBest` class of the Python module `sklearn.feature_selection`. Equation (1) allows us to obtain these scores.

$$I(x; y) = \sum_{i=1}^n \sum_{j=1}^n p(x(i), y(j)) \cdot \log \left(\frac{p(x(i), y(j))}{p(x(i)) \cdot p(y(j))} \right), \quad (1)$$

- where $p(x(i), y(j))$ is the joint probability density function of x and y , and $p(x(i))$ and $p(y(j))$ are the marginal density functions. In Figure 4a, we show the fifteen most relevant features obtained with this technique.

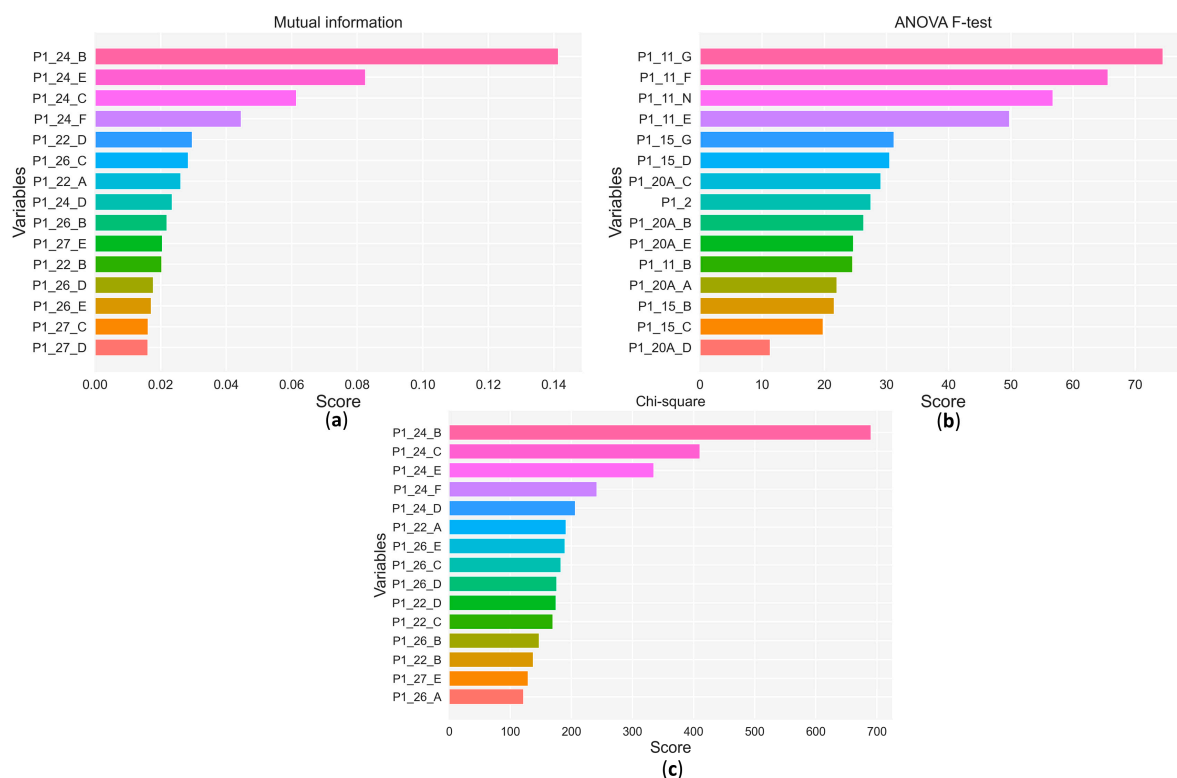


Figure 4. Feature selection using filtering methods: (a) mutual information scores for predicting teachers' life satisfaction; (b) ANOVA scores for predicting teachers' life satisfaction; (c) chi-square scores for predicting teachers' life satisfaction.

- Analysis of variance (ANOVA F-test) is used to compare the means of different groups and determine whether at least one of the means is significantly different from the others [114]. In the context of feature selection, it is used to assess the relevance of a feature in terms of predicting the target variable [115,116]. In this study, since our target variable is categorical, we use this technique to select numerical features. To do so, we employ the `SelectKBest` class, with the parameters `score_func = f_classif` and `k`

= 'all' from the Python module `sklearn.feature_selection`. Equation (2) allows us to obtain the score of this technique.

$$F = \frac{MSB}{MSW},$$

$$MSB = \frac{n_i \sum_{i=1}^k (\bar{y}_i - \bar{y})^2}{(k-1)},$$

$$MSW = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_{ij} - \bar{y}_i)^2}{(N-k)},$$
(2)

- where MSB is the mean of squares between groups, n_i is the number of samples in group i , \bar{y}_i is the mean of group i , \bar{y} is the overall mean of all groups, and k is the number of groups.
- MSW is the mean of the squares within groups, \bar{y}_{ij} is the value of sample j in group i , and N is the total number of samples.
- We show the ANOVA F-test filter scores for the prediction of teachers' life satisfaction in Figure 4b.
- Chi-square analysis is used to determine whether there is a significant association between two categorical variables [114]. In feature selection, this test is used to assess the relevance of a feature in predicting a target variable [117]. In this study, since our target variable is categorical, we applied this technique to select categorical features. We use the `SelectKBest` class, with parameters `score_func = chi2` and `k = 'all'`, from the Python module `sklearn.feature_selection`. Equation (3) shows how the score is calculated for each feature using the following filter.

$$X = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i},$$
(3)

- where O_i is the observer frequency and E_i is the expected frequency.
- In Figure 4c, we show the scores obtained with this filter.
- Spearman correlation coefficient is a nonparametric measure that evaluates the monotonic relationship between two variables based on the ranges of the data rather than their exact values. It is useful in feature selection in ML to evaluate ordinal or monotonic dependencies between features and the target variable, without requiring assumptions about the distribution of the data [118,119]. We use the `spearmanr()` function of the Python module `scipy.stats` to determine the value of the coefficients. Equation (4) allows the calculation of these values.

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}, d = R_x - R_y,$$
(4)

- where R_x and R_y are the ranks of the x e y variables, respectively.
- In Figure 5, we show the Spearman correlation matrix between the fifteen most important variables and the variable "teacher life satisfaction".

3.3.2. Feature Selection by Integrated Methods

The integrated method is more efficient and practical in terms of computational resources compared to the wrapper method, while maintaining a similar performance level. This is because it avoids the need to repeatedly run the classifier and examine each feature subset separately [119]. These methods incorporate feature selection as part of the training process [106,120]. In this research, we employ the methods embedded in the classification and regression trees (CART), Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost algorithms. In Figure 6, we show the results of the application of these methods.

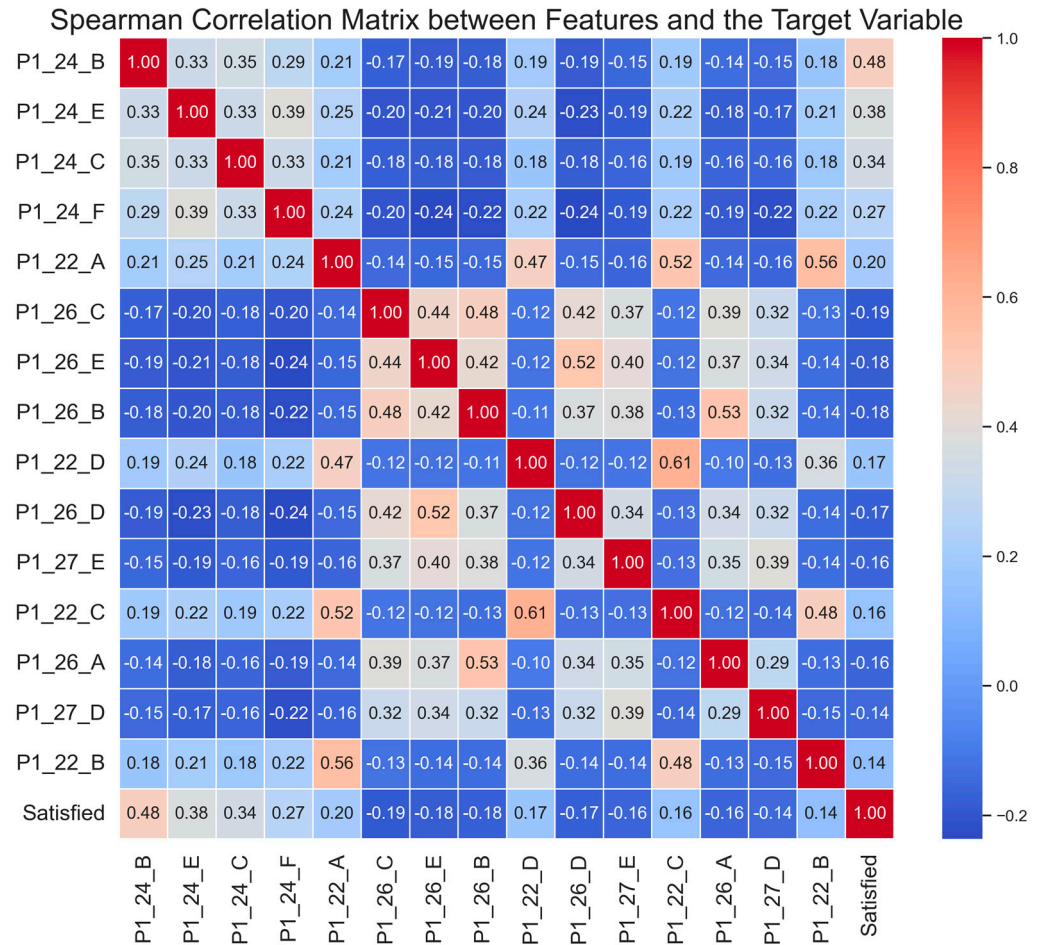


Figure 5. Feature selection performed using Spearman’s Correlation Coefficient to predict teachers’ life satisfaction.

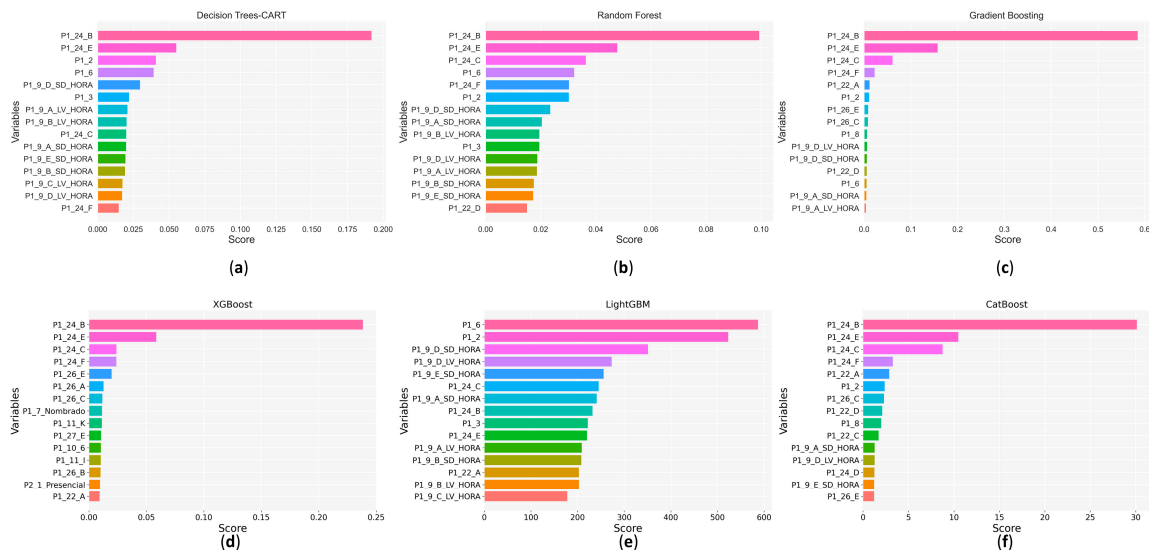


Figure 6. Feature selection using integrated methods to predict teachers’ life satisfaction: (a) CART feature importance; (b) Random Forest feature importance; (c) Gradient Boosting feature importance; (d) XGBoost feature importance; (e) LightGBM feature importance; (f) CatBoost feature importance.

3.3.3. Feature Selection through Ensemble of Methods

Our approach proposes the combination of filter methods and integrated or embedded methods for feature selection in predicting teachers' life satisfaction. Figure 7 shows the algorithms employed in our study.

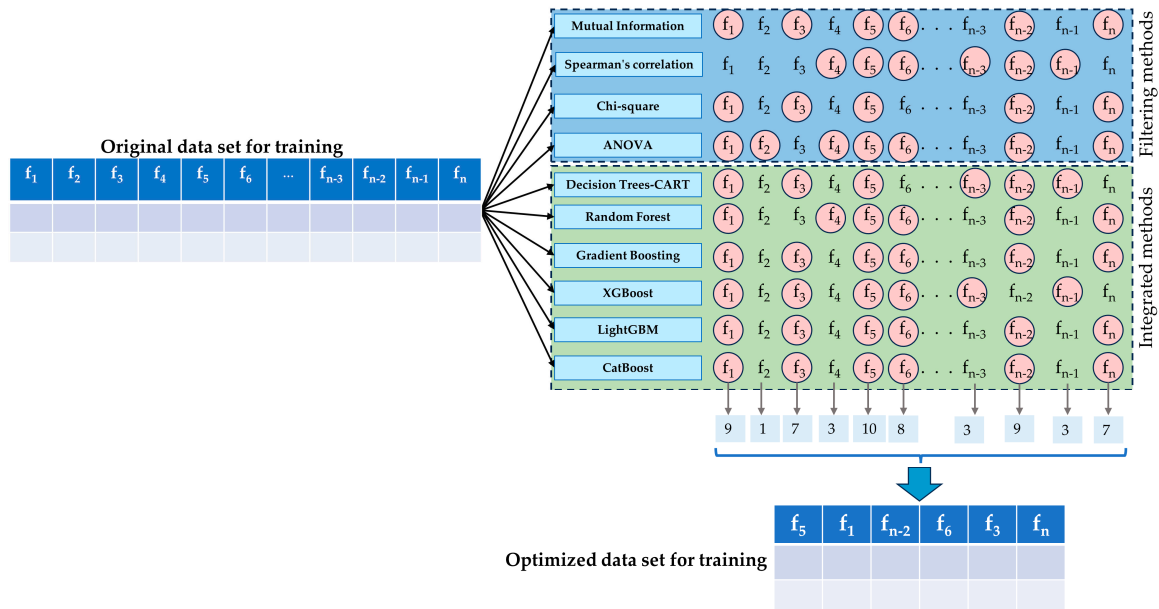


Figure 7. Assembly of methods proposed for selection of characteristics to predict teachers' life satisfaction.

3.3.4. Subset of Data with Characteristics Most Relevant to Satisfaction with Teachers' Lives

To construct the subset of data with the most relevant features in terms of predicting teachers' life satisfaction, we opted for the majority voting approach. The majority voting approach consists in multiple feature selection algorithms being applied to the dataset, where each one "votes" for the features it considers important [76]. The features selected by the majority of algorithms are retained in the final feature set. This approach seeks to combine the strength of several methods to obtain a more robust and reliable selection.

In this study, we included variables selected by at least 2 algorithms, resulting in a total of 28 features. To determine the optimal number of features, we trained a logistic regression model with default hyperparameters. This algorithm was chosen because it was not considered during feature selection or at the model training and evaluation stage, making it neutral in our study. We evaluated the contribution of each variable using six metrics: balanced accuracy, F1-score, precision, recall, Cohen's kappa coefficient, and Jaccard Score.

In Figure 8, we show that, by using the first 10 variables, the metrics are stabilized. This suggests that these variables form the subset of data with the most relevant characteristics.

Table 2 provides an overview of the subset of data, that includes the 10 most relevant features for predicting teachers' life satisfaction. This dataset consists of 18,911 instances and 11 columns: 10 of them represent predictor variables and 1 represents the target variable ("satisfied"). In the "satisfied" variable, 0 represents "not satisfied", 1 represents "satisfied", and 2 represents "very satisfied". From this dataset, 15,128 instances correspond to the training set and 3783 relate to the test set. A detailed description of each of these features is found in Table A1.

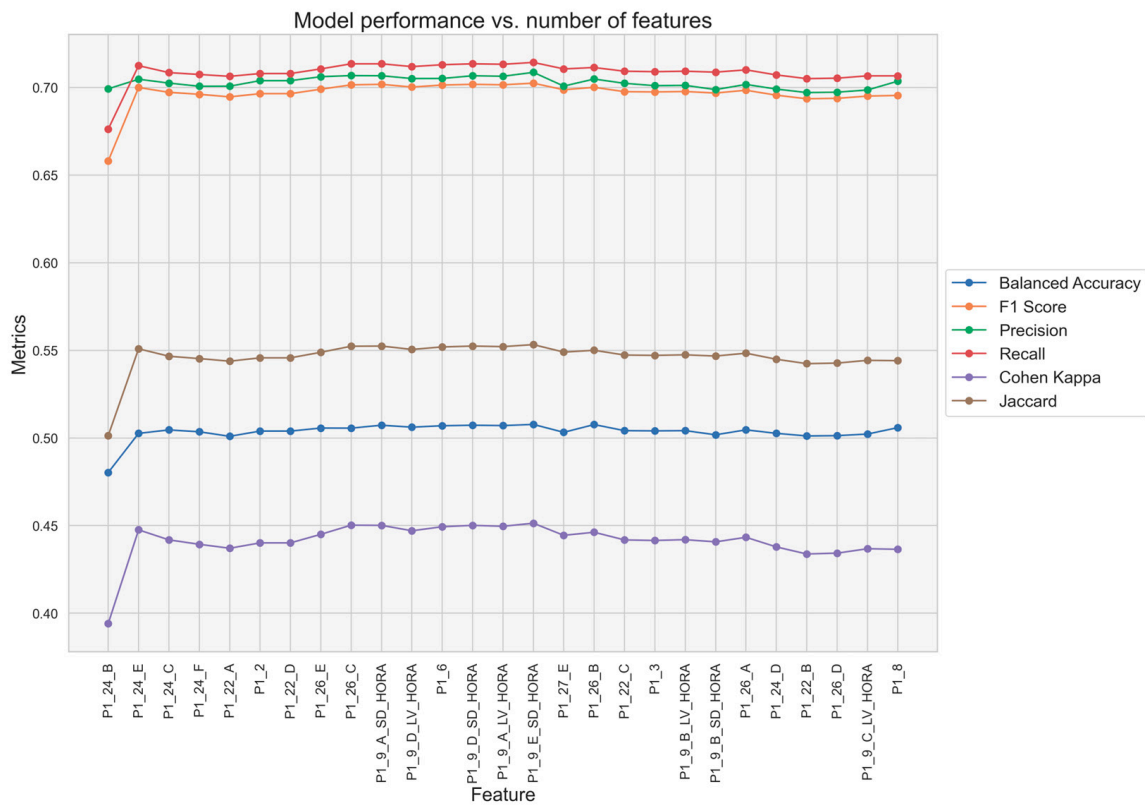


Figure 8. Evolution of metrics versus important variables in predicting teachers’ life satisfaction.

Table 2. Subset of data with the most relevant features for predicting teachers’ life satisfaction.

P1_24_B	P1_24_E	P1_24_C	P1_24_F	P1_22_A	P1_2	P1_22_D	P1_26_E	P1_26_C	P1_9_A_SD_HORA	Satisfied
3.0	3.0	3.0	3.0	2.0	1.022093	3.0	1.0	1.0	-0.550070	2
2.0	2.0	2.0	1.0	1.0	1.869359	1.0	1.0	1.0	-0.970063	1
2.0	2.0	2.0	1.0	2.0	1.869359	2.0	2.0	1.0	1.969885	1
...
3.0	3.0	3.0	3.0	2.0	-0.460624	3.0	1.0	1.0	1.129900	2
3.0	3.0	2.0	3.0	2.0	-0.460624	3.0	1.0	1.0	0.289915	2
3.0	3.0	2.0	2.0	2.0	0.068918	2.0	1.0	1.0	-0.550070	2

3.4. Training and Model Evaluation

We performed our experiments in Jupyter Notebook using the following libraries: we applied Scikit-learn for data preprocessing, feature selection, and model training and evaluation; we applied R’s matplotlib and ggplot2 for data visualization; and we applied numpy for numerical computations. All these libraries were installed in the Anaconda Navigator 2.6.0 suite, an open-source environment and package manager for Python and R. Anaconda Navigator was installed on a computer with the following features: an Intel(R) Core (TM) i5-12450H 12th generation 2.50 GHz processor, 40.0 GB RAM, and a Windows 11 Home 64-bit operating system.

3.4.1. Training and Hyperparameters Tuning

In this research, we trained models using a variety of algorithms, including Random Forest, XGBoost, Gradient Boosting, Decision Trees—CART, CatBoost, LightGBM, Support Vector Machine, and Multilayer Perceptron. We performed the fine-tuning of hyperparameters to find the most optimal ones, because these maximize the performance of the ML models [121].

There are two common approaches to hyperparameter fine-tuning: Grid Search and Random Search [122,123]. Grid Search performs an exhaustive exploration of all possible combinations within a predefined hyperparameter space, evaluating each configuration to

identify the most optimal one [124,125]. Although this method guarantees the evaluation of all combinations, it can be computationally expensive, especially when the number of hyperparameters is large and the range of values is wide. On the other hand, Random Search selects random combinations of hyperparameters within the defined search space [121]. This approach allows for a wider and more efficient exploration of the space, with a smaller number of evaluations [126,127]. In our research, we use Random Search with the stratified cross-validation (StratifiedKFold) of 100 different random combinations of hyperparameters, with $K = 10$. This means that the data are divided into 10 subsets (folds), each maintaining the same class distribution as the original dataset, which is ideal for classification problems with unbalanced classes [122]. The model is trained and evaluated 10 times, using 9 folds for training and 1 for validation, thus obtaining 10 accuracy scores, the average of which is used as the final score. Figure 9 shows this process graphically.

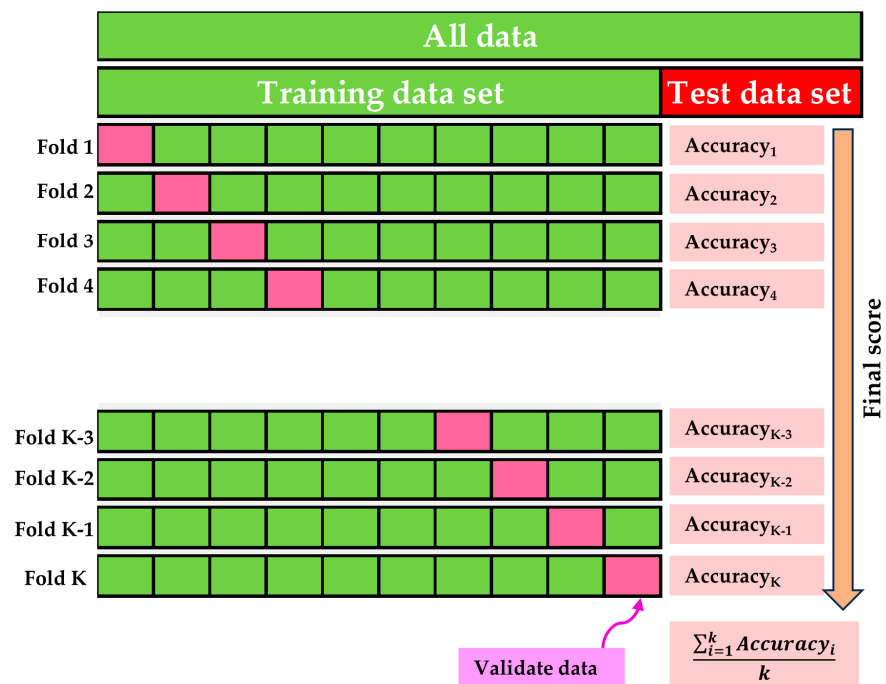


Figure 9. Cross-validation process.

Table 3 presents the hyperparameters for each model, including the search space, description, default values, and optimized values.

Table 3. Hyperparameter fine-tuning.

Model	Hyperparameter	Search Space	Description	Default Values	Optimal Values
Random Forest	n_estimators	[10:100] step 1	Number of trees	100	85
	criterion	["gini", "entropy"]	Criteria for evaluating divisions	"gini"	"entropy"
	max_depth	[2:20] step 1	Maximum depth	None	None
	min_samples_split	[2:10] step 1	Minimum number of samples to split node	2	2
	min_samples_leaf	[1:10] step 1	Minimum samples to be a leaf node	1	1
	max_features	["auto", "sqrt", "log2", None]	Number of characteristics to consider for the best division	"sqrt"	"sqrt"
	bootstrap	[True, False]	Method for sampling input data	True	True
XGBoost	n_estimators	[10, 17, 25, 33, 41, 48, 56, 64, 72, 80]	Number of trees	None	80
	max_depth	[3, 5, 7]	Maximum Depth	None	3
	learning_rate	[0.01:0.1] step 0.03	Learning rate	None	0.01
	subsample	[0.6:0.9] step 0.1	Proportion of samples to train each tree	None	0.8
	colsample_bytree	[0.6:0.9] step 0.1	Proportion of features per tree	None	0.8

Table 3. Cont.

Model	Hyperparameter	Search Space	Description	Default Values	Optimal Values
Gradient Boosting	loss	["log_loss"]	Loss function	"log_loss"	"log_loss"
	learning_rate	[0.001, 0.005, 0.01, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2]	Learning rate	0.1	0.025
	min_samples_split	[500:595] step 5 + [601:696] step 5 + [702:797] step 5 + [803:898] step 5 + [904:1000] step 5	Minimum samples to split a node	2	606
	min_samples_leaf	[20, 28, 37, 46, 55, 64, 73, 82, 91, 100]	Minimal samples in a leaf node	1	100
	max_depth	[2:10] step 1	Maximum tree depth	3	8
	max_features	["log2", "sqrt"]	Number of characteristics to consider for the best division	None	"sqrt"
	criterion	["friedman_mse", "squared_error"]	Criteria for evaluating divisions	"friedman_mse"	"squared_error"
	subsample n_estimators	[0.5, 0.618, 0.8, 0.85, 0.9, 0.95, 1.0] [100:1000] step 100	Proportion of samples to train each tree Number of sequential trees	1.0 100	0.618 200
Decision Trees-CART	max_depth	[10, 20, 30, 40, 50, None]	Maximum tree depth	None	None
	criterion	["gini", "entropy"]	Criteria for measuring the quality of a division	"gini"	"entropy"
	min_samples_split	[2, 3, 4, 5, 7, 10, 15]	Minimum samples to split a node	2	2
	min_samples_leaf	[1, 2, 3, 4, 5, 7]	Minimal samples in a leaf node	1	3
	max_features	["sqrt", "log2"]	Maximum number of characteristics to be considered for a division	None	"sqrt"
CatBoost	iterations	[100:500] step 100	Number of iterations (trees)	1000	400
	depth	[3: 10] step 1	Maximum tree depth	6	10
	learning_rate	[0.01, 0.05, 0.1, 0.2]	Learning rate	0.093	0.2
	l2_leaf_reg	[1:9] step 2	L2 regularization in leaf	3.0	1
	border_count	[32, 50, 100, 200]	Number of division limits in numerical characteristics	254	32
	bagging_temperature	[0.5, 1, 2, 3]	Controls the intensity of random sampling	1.0	3
	random_strength	[1, 2, 5, 10]	Intensity of random noise to handle equal predictions	1.0	5
	one_hot_max_size	[2, 10, 20]	Maximum size to use one-hot encoding	2 ¹⁵	2
LightGBM	num_leaves	[20:140] step 10	Maximum number of leaves per tree	31	80
	max_depth	[3, 5, 7, 9, 11, 13]	Maximum depth	-1	9
	learning_rate	[0.0001, 0.001, 0.01, 0.1, 1.0]	Learning rate	0.1	0.1
	n_estimators	[100, 300, 500, 700, 900]	Number of trees	100	300
	min_child_samples	[5, 15, 25, 35, 45]	Minimum samples in leaf nodes	20	35
	subsample	[0.6, 0.7, 0.8, 0.9, 1.0]	Proportion of data to train each tree	1.0	0.7
	colsample_bytree	[0.6, 0.7, 0.8, 0.9, 1.0]	Proportion of characteristics per tree	1.0	0.8
	reg_alpha	[1.0×10^{-4} , 1.78×10^{-3} , 3.16×10^{-2} , 5.62×10^{-1} , 1.0×10^1]	Regularization L1	0.0	3.16×10^{-2}
	reg_lambda	[1.0×10^{-4} , 1.78×10^{-3} , 3.16×10^{-2} , 5.62×10^{-1} , 1.0×10^1]	Regularization L2	0.0	1.78×10^{-3}
	min_split_gain	[0.0, 0.25, 0.5, 0.75, 1.0]	Minimum gain for splitting a node	0.0	0.0
scale_pos_weight	[1, 10, 25, 50, 75, 99]	Balancing unbalanced classes	1.0	10	
Support Vector Machine	C	[0.1, 1, 10, 100, 1000]	Regularization parameter	1.0	0.1
	gamma	[1, 0.1, 0.01, 0.001, 0.0001]	Kernel coefficient	"scale"	1
	kernel	["linear", "rbf"]	Kernel function	"rbf"	"linear"
Multilayer Perceptron	hidden_layer_sizes	[50, 100, 150]	Number of neurons in the hidden layer	100	150
	activation	["tanh", "relu"]	Activation function	"relu"	"relu"
	solver	["adam", "sgd"]	Optimization method	"adam"	"adam"
	alpha	[0.0001, 0.001, 0.01]	Adjustment parameter	0.0001	0.0001
	learning_rate	["constant", "adaptive"]	Learning rate	"constant"	"constant"
	max_iter	Entero aleatorio entre 100 y 1000	Number of training iterations	200	848

3.4.2. Model Evaluation

In order to evaluate the performance of the algorithms used in our study, we applied Tukey's test [126]. This test is fundamental in comparative analyses such as the present one, where we evaluate multiple algorithms in terms of various performance metrics. Tukey's test allows us to accurately identify which groups of algorithms show significant differences in their performance, ensuring a robust comparison and controlling for type I errors [124]. The data used in this test were generated using the 100-fold K-Fold stratified cross-validation technique.

In addition to Tukey's test, we evaluated the performance of the algorithms on the test dataset, taking as references the confusion matrix and the metrics derived from it, such as accuracy, balanced accuracy, recall, precision, F1-Score, Cohen's kappa coefficient, and Jaccard Score. Each of these metrics is detailed below.

A confusion matrix is a tool that allows the visualization of the distribution of false predictions. It provides a detailed analysis of the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) [128].

Figure 10 illustrates a confusion matrix for a classification problem with three classes (A, B and C). The elements AA, BB, and CC represent the true positives of classes A, B and C, respectively, meaning that the samples were correctly classified. However, AB symbolizes the samples in class A that were misclassified as class B. The class A false negatives (FN_A) are determined by adding AB and AC ($FN_A = AB + AC$), which correspond to the class A samples that were incorrectly classified as class B or C. For any class in a row, false negatives are calculated by summing the errors in that row. The false positives for a predicted class are obtained by summing all the errors in the corresponding column. Thus, the false positive of class A (FP_A) is the sum of BA and CA ($FP_A = BA + CA$).

		Predicted Class			FN
		A	B	C	
True Class	A	AA	AB	AC	AB+AC
	B	BA	BB	BC	BA+BC
	C	CA	CB	CC	CA+CB
FP		BA+CA	AB+CB	AC+BC	

Figure 10. Confusion matrix for multiclass classification problems.

Accuracy is defined as the proportion of correct predictions made by the model with respect to the total number of predictions. We define it formally by means of Equation (5).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Balanced accuracy is a metric that considers both the sensitivity or recall of each class, and which is especially useful in classification problems with unbalanced classes. Equation (6) formally defines this metric.

$$Balanced\ Accuracy = \frac{1}{N} \sum_{i=1}^N Recall_i, \quad (6)$$

where $Recall_i$ is the recall for class i .

Recall, also known as sensitivity or True Positive Rate, is a metric that measures the ability of a model to correctly identify positive observations. Equation (7) allows the calculation of this metric.

$$Recall_i = \frac{TP_i}{TP_i + FN_i}, \quad (7)$$

where TP_i indicates the true positives for class i , i.e., the number of samples of class i that were correctly classified, and FN_i indicates the false negatives for class i , i.e., the number of samples of class i that were incorrectly classified as belonging to any other class.

Precision is defined as the quotient between the number of true positives and the total sum of positives. The formal definition of this metric is presented in Equation (8).

$$Precision_i = \frac{TP_i}{TP_i + FP_i}, \quad (8)$$

where TP_i indicates the true positives for class i and FP_i indicates false positives for class i .

F1-Score combines precision and recall into a single metric, being useful for evaluating model performance, especially in unbalanced datasets. In this study, we use Weighted

F1-Score; this approach calculates the F1-Score for each class and weights it according to the number of instances of each class. Equation (9) allows us to calculate this value.

$$F1 - Score_i = 2 \left(\frac{Precision_i \times Recall_i}{Precision_i + Recall_i} \right),$$

$$Weighted\ F1 - Score = \sum_{i=1}^N \left(\frac{n_i}{n} \times F1 - Score_i \right), \quad (9)$$

where $Precision_i$ is the precision of class i , and $Recall_i$ is the recall of class i .

Cohen's kappa coefficient evaluates the agreement between two evaluators in classification problems [129]. A value of 1 represents perfect agreement, while a kappa of 0 indicates an agreement no better than what would be expected by chance.

In the context of ML, Cohen's kappa coefficient is used to measure the agreement between the predictions of a model and the actual labels in classification problems. Equation (10) shows the calculation of this metric.

$$kappa = \frac{p_o - p_e}{1 - p_e}, \quad (10)$$

where p_o is the proportion of observations in which the evaluators agree and p_e is the proportion of expected agreements by chance.

Jaccard Score, also known as the Jaccard index, is a metric that evaluates how similar two datasets are. In the context of multiclass classification, it is determined by calculating the ratio of the intersection size to the union size of the label sets. We show the calculation of this metric in Equation (11).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (11)$$

where A represents the predicted labels and B the reference labels.

4. Results

We present the results based on metrics such as confusion matrix, accuracy, balanced accuracy, recall, F1-Score, precision, Cohen kappa coefficient, and Jaccard Score, obtained through 100-fold stratified K-Fold cross-validation of the training dataset. Additionally, we applied the Tukey statistical test to compare these metrics across the algorithms. Finally, we validated the aforementioned metrics on the test dataset.

In Figure 11, we present the most relevant features in terms of predicting the life satisfaction of Peruvian teachers. The variables P1_24_B, P1_24_C, P1_24_E, P1_24_F, P1_22_A, P1_2, P1_22_D, P1_26_C, P1_26_E, P1_9_D_IV_HORA, and P1_9_A_SD_HORA are used in the majority of methods, with their high selection frequency highlighting their importance. For a more detailed description of each variable, we present Table A1, which provides a data dictionary, with definitions and codes corresponding to each feature selected by the ensemble of methods.

Table 4 shows the average performance of the algorithms on the training dataset, using 100-fold stratified K-Fold cross-validation. Notably, the CatBoost and LightGBM models achieved the highest values across all evaluated metrics. Specifically, CatBoost reached an accuracy, balanced accuracy, and recall of 0.824, with a precision and F1-Score of 0.82. On the other hand, LightGBM obtained an accuracy, balanced accuracy, recall, precision, and F1-Score of 0.81, indicating high consistency and the ability to balance classes effectively. Both models also excel in Cohen kappa (CatBoost: 0.74, LightGBM: 0.72) and Jaccard Score (CatBoost: 0.71, LightGBM: 0.70), reflecting high agreement and a good intersection over union for the predicted labels with the true ones. Random Forest also shows a strong performance, with a precision of 0.79. However, the Gradient Boosting, Support Vector Machine, and XGBoost models show significantly lower performances, with precisions of 0.68, 0.62, and 0.63, respectively.

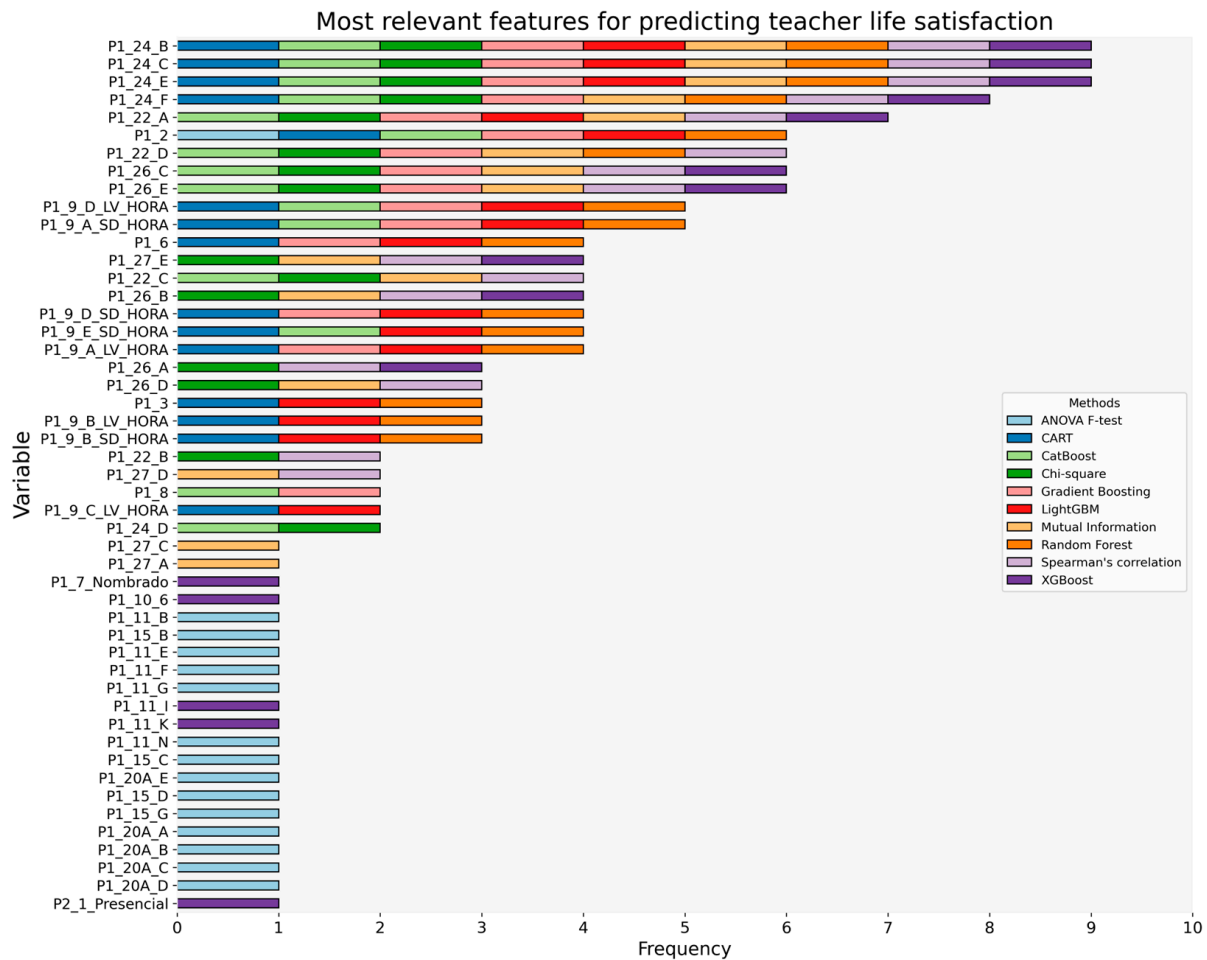


Figure 11. Most important features for predicting the life satisfaction of Peruvian teachers.

Table 4. Summary statistics of the metrics obtained on the training dataset.

Model	Accuracy	Balanced Accuracy	Recall	Precision	F1 Score	Cohen Kappa Coefficient	Jaccard Score
CatBoost	0.824 ± 0.026	0.824 ± 0.026	0.824 ± 0.026	0.823 ± 0.027	0.822 ± 0.027	0.737 ± 0.039	0.714 ± 0.036
CART	0.762 ± 0.026	0.762 ± 0.026	0.762 ± 0.026	0.756 ± 0.028	0.755 ± 0.028	0.642 ± 0.039	0.622 ± 0.034
Gradient Boosting	0.677 ± 0.029	0.677 ± 0.029	0.677 ± 0.029	0.677 ± 0.029	0.676 ± 0.029	0.515 ± 0.043	0.516 ± 0.033
LightGBM	0.814 ± 0.024	0.814 ± 0.024	0.814 ± 0.024	0.811 ± 0.025	0.811 ± 0.025	0.721 ± 0.036	0.698 ± 0.033
MLP classifier	0.735 ± 0.026	0.735 ± 0.026	0.735 ± 0.026	0.735 ± 0.027	0.732 ± 0.026	0.603 ± 0.039	0.586 ± 0.032
Random Forest	0.791 ± 0.024	0.791 ± 0.024	0.791 ± 0.024	0.787 ± 0.025	0.787 ± 0.025	0.687 ± 0.036	0.661 ± 0.032
SVM	0.615 ± 0.032	0.615 ± 0.032	0.615 ± 0.032	0.644 ± 0.031	0.619 ± 0.031	0.422 ± 0.048	0.451 ± 0.033
XGBoost	0.633 ± 0.032	0.633 ± 0.032	0.633 ± 0.032	0.634 ± 0.032	0.631 ± 0.032	0.449 ± 0.048	0.466 ± 0.034

Figure 12 graphically shows the distribution of the evaluated metrics. The graph highlights that the CatBoost, LightGBM and Random Forest algorithms show higher medians in all the evaluated metrics compared to the other algorithms. Although some outliers are observed in these models, their performance remains robust, as the inter-quartile variability, indicated by the length of the boxes, is lower compared to the other algorithms.

Table 5 shows the results of the one-factor ANOVA test applied to the balanced accuracy, Sensitivity, F1-Score and Cohen’s kappa coefficient metrics. These metrics, which are particularly useful for dealing with unbalanced datasets. The last column reflects *p*-value = 0.00, indicating a *p*-value < 0.05 for all metrics evaluated. This suggests a significant difference in the performance of at least two of the evaluated algorithms.

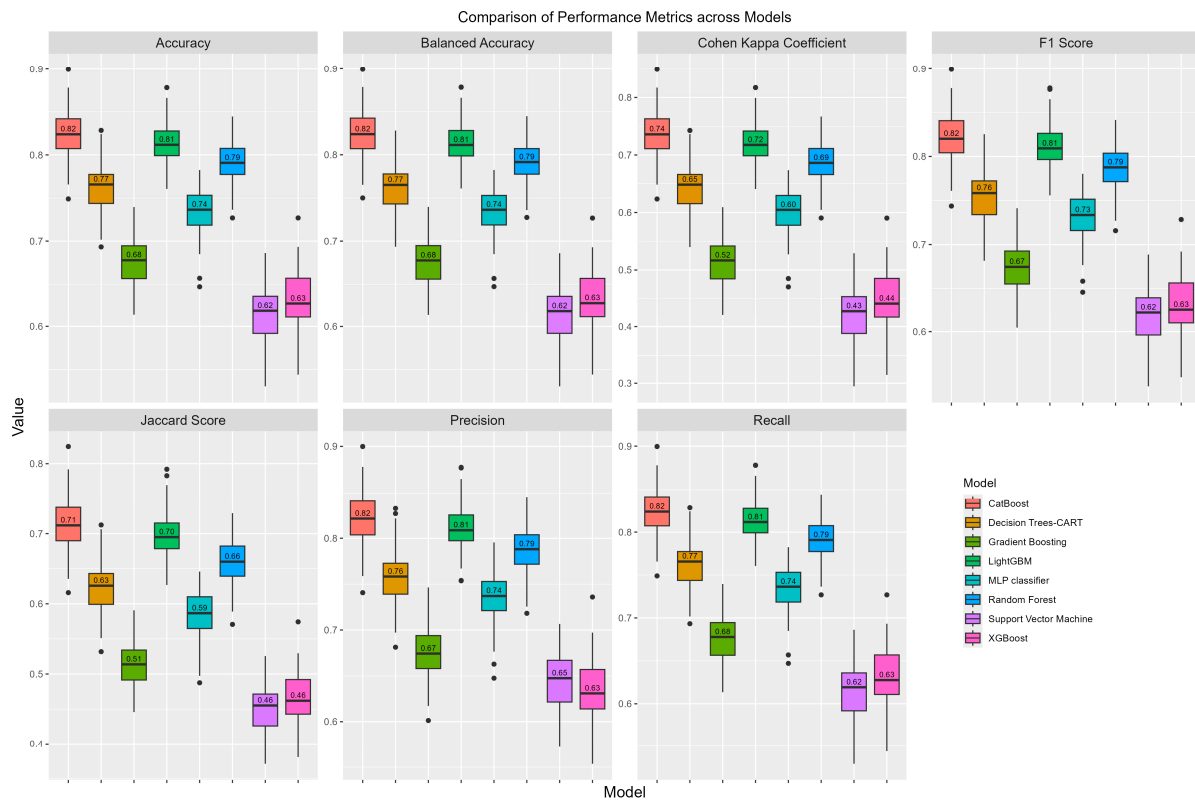


Figure 12. Distribution of evaluation metrics in training dataset.

Table 5. ANOVA test on the metrics evaluated.

	Sum of Squares	df	Mean Square	F	Sig.	
Balanced accuracy	Between groups	4.633	7	0.662	880.466	0.000
	Within groups	0.595	792	0.001		
	Total	5.229	799			
Sensitivity	Between groups	4.633	7	0.662	881.058	0.000
	Within groups	0.595	792	0.001		
	Total	5.228	799			
F1 Score	Between groups	4.414	7	0.631	804.628	0.000
	Within groups	0.621	792	0.001		
	Total	5.035	799			
Cohen kappa coefficient	Between groups	10.425	7	1.489	881.251	0.000
	Within groups	1.338	792	0.002		
	Total	11.763	799			

Table 6 presents the results of the Tukey test for the metrics of balanced accuracy, sensitivity, F1-Score, and Cohen kappa coefficient. With a significance level of 0.05, according to the Tukey test, we can state that group 7, composed of the CatBoost and LightGBM algorithms, has a superior average performance compared to the other algorithms evaluated. The Random Forest algorithm, which is used in group 6, closely follows in terms of performance. Additionally, the Support Vector Machine algorithm shows the lowest average performance across all evaluated metrics.

In Figure 13, we present the confusion matrices of the constructed models, where the classes represent the following meanings: 0—“not satisfied”; 1—“satisfied”; and 2—“very satisfied”. The XGBoost, Random Forest, and LightGBM algorithms excel in classifying the “very satisfied” class, with high true positive values (XGBoost: 1482, Random Forest: 1418, LightGBM: 1409) indicating high precision in identifying the most satisfied teachers. On the other hand, CatBoost and LightGBM are effective in classifying the “satisfied” class,

with high true positive values (CatBoost: 1109, LightGBM: 1108) demonstrating a greater ability to identify satisfied teachers. Regarding the “not satisfied” class, the XGBoost and Support Vector Machine algorithms stand out with 98 true positives each, indicating their effectiveness in identifying dissatisfied teachers.

Table 6. Results of the Tukey test for the metrics evaluated.

HSD Tukey ^a		Balanced Accuracy						HSD Tukey ^a		Sensitivity								
Model	N	1	2	Subset for Alpha= 0.05			6	7	Model	N	1	2	Subset for Alpha = 0.05			6	7	
				3	4	5							3	4	5			
Support Vector Machine	100	0.615							Support Vector Machine	100	0.615							
XGBoost	100		0.633						XGBoost	100		0.633						
Gradient Boosting	100			0.677					Gradient Boosting	100			0.677					
MLP Classifier	100				0.735				MLP Classifier	100				0.735				
Decision Trees—CART	100					0.762			Decision Trees—CART	100					0.762			
Random Forest	100						0.791		Random Forest	100						0.791		
LightGBM	100							0.814	LightGBM	100							0.814	
CatBoost	100							0.824	CatBoost	100							0.824	
Sig.		1.000	1.000	1.000	1.000	1.000	1.000	0.117	Sig.		1.000	1.000	1.000	1.000	1.000	1.000	0.117	

HSD Tukey ^a		F1 Score						HSD Tukey ^a		Cohen Kappa Coefficient								
Model	N	1	2	Subset for Alpha = 0.05			6	7	Model	N	1	2	Subset for Alpha = 0.05			6	7	
				3	4	5							3	4	5			
Support Vector Machine	100	0.619							Support Vector Machine	100	0.422							
XGBoost	100		0.631						XGBoost	100		0.449						
Gradient Boosting	100			0.676					Gradient Boosting	100			0.515					
MLP Classifier	100				0.732				MLP Classifier	100				0.603				
Decision Trees—CART	100					0.755			Decision Trees—CART	100					0.642			
Random Forest	100						0.787		Random Forest	100						0.687		
LightGBM	100							0.811	LightGBM	100							0.721	
CatBoost	100							0.822	CatBoost	100							0.737	
Sig.		1.000	1.000	1.000	1.000	1.000	1.000	0.125	Sig.		1.000	1.000	1.000	1.000	1.000	1.000	0.117	

The means for the groups in the homogeneous subsets are displayed. ^a using the sample size of the harmonic mean = 100,000.

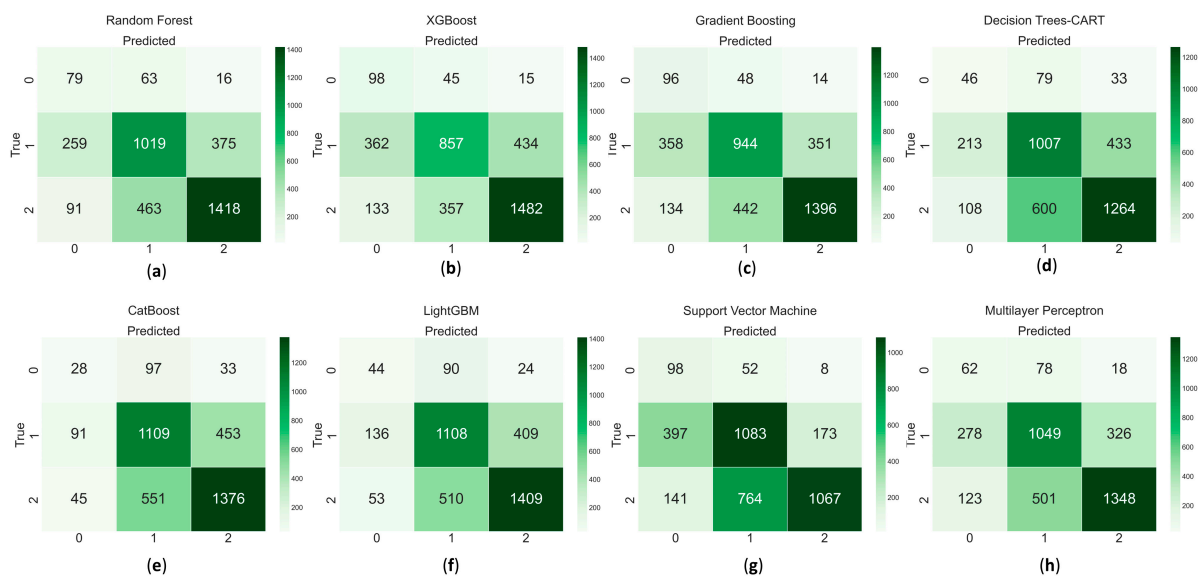


Figure 13. Confusion matrices of the evaluated models: (a) Random Forest; (b) XGBoost; (c) Gradient Boosting; (d) Decision Trees—CART; (e) CatBoost; (f) LightGBM; (g) Support Vector Machine; (h) Multilayer Perceptron.

With respect to false negatives, XGBoost shows the lowest number of false negatives (60 teachers) for class 0, meaning “not at all satisfied”. This suggests its effectiveness in minimizing errors in terms of classifying dissatisfied teachers. For class 1, “satisfied”, CatBoost shows the lowest number of false negatives (544 teachers), demonstrating its accuracy in correctly identifying satisfied teachers. For class 2, “very satisfied”, XGBoost

again shows the lowest value (490 teachers) of false negatives, indicating its robustness in identifying the most satisfied teachers.

In Figure 14, we show that the XGBoost and Gradient Boosting models excel in terms of the recall metric, both showing values of 0.63. These results suggest that these models are effective in identifying true positives, which is crucial in an educational context in order to ensure that both satisfied and dissatisfied teachers are correctly identified. The ability to accurately identify these instances can better inform decisions regarding interventions and support for teachers. However, the highest precision is observed in the LightGBM and Random Forest models, both with a value of 0.55, indicating that these models have a lower false positive rate.

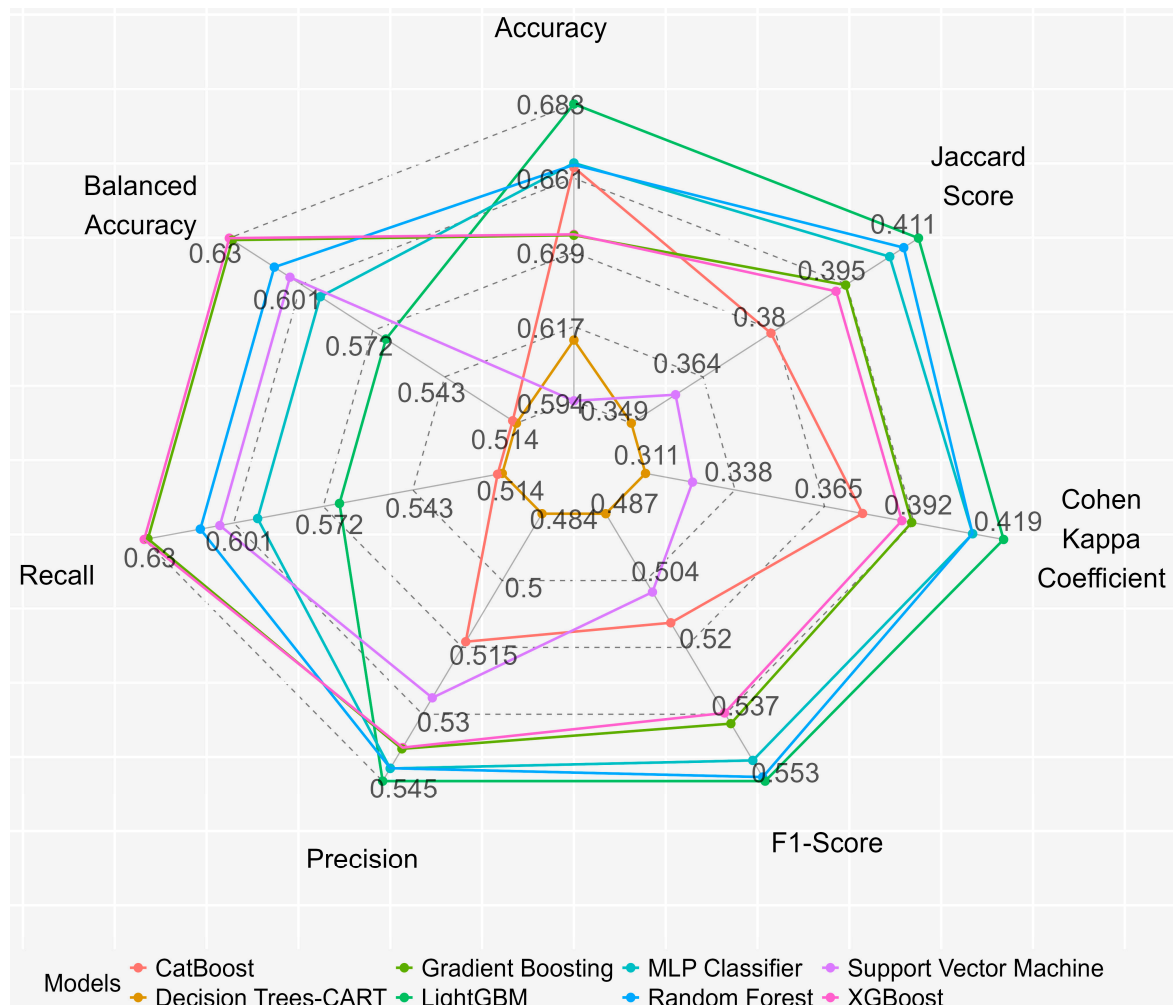


Figure 14. Results of the metrics evaluated in the test dataset.

Regarding the F1-Score metric, the highest value is 0.55, achieved by the LightGBM and Random Forest models. This indicates that these models display a good balance between precision and recall, which is crucial for practical applications in education where both aspects are important. Additionally, the Cohen kappa coefficient, which measures the agreement between predicted and actual labels, is highest in LightGBM (0.42) and Random Forest (0.41) assessments, reinforcing the idea that these models have greater consistency in their predictions.

Regarding the F1-Score metric, the LightGBM and Random Forest models have the highest values at 0.55, indicating a good balance between precision and recall. Additionally, the Cohen kappa coefficient, which measures the agreement between predicted and actual

labels, is higher in LightGBM (0.42) and Random Forest (0.41), reinforcing the consistency of these predictions.

Finally, in the Jaccard Score metric, the LightGBM and Random Forest models achieve the highest values at 0.41. These results indicate that LightGBM has a greater ability to correctly match predicted labels with true labels compared to the other models. A higher Jaccard Score reflects better precision and fewer errors.

5. Discussion

In our study, we found that teachers' personal well-being, represented by the variables "Satisfaction with Health" (P1_24_B) and "Satisfaction with Living Conditions for the Family" (P1_24_C), is crucial to their overall life satisfaction. Health is a fundamental component of well-being, as good health allows for greater participation in daily and work activities, which positively contributes to life satisfaction [16]. Moreover, satisfaction with the living conditions that teachers can provide for their children and family suggests that the ability to ensure a suitable living environment is a significant source of personal satisfaction, aligning with the findings reported by [130]. This implies that teachers who are able to care well for their families tend to feel more satisfied with their lives, highlighting the importance of family support and stability in personal well-being [131].

On the other hand, we found that job satisfaction, represented by the variables "Satisfaction with Work in the Educational Institution" (P1_24_E) and "Satisfaction with Conditions for Performing Teaching Duties" (P1_24_F), was relevant in predicting life satisfaction. These findings confirm that job satisfaction is closely linked to overall well-being and work motivation, supporting the results of previous studies [11,132]. This connection highlights the importance of a favorable work environment and adequate teaching conditions in the overall well-being of educators.

In addition to these variables, the ability to reflect on pedagogical outcomes (P1_26_C) and participation in continuous learning processes (P1_26_E) proved essential for teachers' life satisfaction. A teacher who can reflect on their results is able to learn from their experiences and continually improve their practice [133,134]. Similarly, a teacher who actively participates in continuous training programs stays updated on the latest trends in education, pedagogy, and didactics, developing new skills and knowledge to enhance their teaching practice, which increases job satisfaction and emotional well-being [135]. Therefore, educational institutions should create spaces and opportunities for teachers to reflect on their classes and share experiences with colleagues, as well as provide access to training courses, workshops, and other professional development activities.

Our study also reveals that trust in educational and governmental institutions, represented by the variable of trust in the Ministry of Education (P1_22_A) and in the UGEL (P1_22_D), plays a crucial role in teachers' life satisfaction. This finding aligns with previous research suggesting that trust in these institutions is positively associated with perceptions of support and job stability, factors that directly influence personal and professional well-being [136–138].

Another important variable identified is age (P1_2). This finding confirms what is reported by the World Happiness Report [139], which states that, in most countries, life satisfaction gradually declines from childhood through adolescence into adulthood. This implies that, as teachers age, they may face greater personal and professional challenges that negatively impact their life satisfaction. Furthermore, it underscores the need to design well-being policies that consider different life stages to more effectively support teachers throughout their careers.

Finally, teachers' life satisfaction is also influenced by the balance between their work and personal life, as reflected in the variables of the number of hours spent on household chores and caregiving during the week (P1_9_D_LV_HORA) and the number of hours spent preparing lessons and handling administrative tasks during the week (P1_9_A_SD_HORA). This suggests that reaching an appropriate balance between these activities may be crucial for teachers' personal well-being. Research indicates that an imbalance between work

and personal responsibilities can lead to increased stress and life dissatisfaction [140,141]. Therefore, it is essential for teachers to effectively split their time between work demands and personal responsibilities to maintain a high level of life satisfaction. The ability to handle these tasks efficiently not only contributes to their personal well-being but can also enhance their professional performance and job satisfaction [142].

Another interesting finding is that the LightGBM and Random Forest algorithms excel in five of the seven metrics evaluated, including precision, F1-Score, Cohen's kappa coefficient, Jaccard Score, and accuracy. These results suggest that these algorithms could be considered for implementing effective interventions in the educational sector. Precision in predicting teachers' life satisfaction is crucial for identifying dissatisfied teachers and designing appropriate solutions. Our findings are consistent with those reported by Pan and Cutumisu [69], who successfully predicted the life satisfaction of secondary school students in the UK and Japan, where the Random Forest algorithm performed better than K-Nearest Neighbors (KNN) model. Similarly, Khan et al. [70] found that Random Forest predicted life satisfaction of Danish citizens with an accuracy of 93.80% and a macro F1 score of 70.60%. Jaiswal et al. [71] also reported that Random Forest was able to predict overall happiness with a precision of 92.27%, the highest among the algorithms evaluated.

In this study, we explored new predictors of life satisfaction among Peruvian teachers using an ensemble approach of feature selection methods and machine learning. According to our literature review, there are no studies addressing this issue with our approach, which adds novelty and interest to our work. However, it is important to note some limitations. First, the data analyzed are from 2020, and although the ENDO 2021 exists, it does not include as many variables as the ENDO 2020. Second, although Grid Search is more thorough in exploring hyperparameters, we did not use it due to its high computational cost, opting instead for Random Search. Future studies are expected to benefit from more powerful computing resources and more up-to-date data.

6. Conclusions

In this study, we identified the most relevant predictors of life satisfaction among basic education teachers in Peru using data from the ENDO-2020. We employed an ensemble approach to feature selection methods, combining filtering techniques such as mutual information, analysis of variance, chi-square, and Spearman's Rank Correlation Coefficient with integrated methods like classification and regression trees, Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost.

Our results show that the most relevant variables in terms of predicting teachers' life satisfaction are satisfaction with their health, their employment at educational institutions, the living conditions they can provide for their family, and the conditions for performing their teaching duties. Other important variables include age, the degree of trust in the Ministry of Education and the Local Management Unit (UGEL), participation in continuous training programs, reflection on the outcomes of their teaching practice, work-life balance, and the number of hours dedicated to lesson preparation and administrative tasks.

In addition, we found that the LightGBM and Random Forest algorithms performed the best on five of the seven metrics evaluated. Using the test dataset, LightGBM achieved an accuracy of 0.55, an F1-Score of 0.55, a Cohen's kappa coefficient of 0.42, and a Jaccard Score of 0.41. Random Forest achieved an accuracy of 0.54, an F1-Score of 0.55, Cohen's kappa coefficient of 0.41 and a Jaccard Score of 0.41. These results suggest that LightGBM and Random Forest are the most suitable models for predicting teachers' life satisfaction.

Our findings provide a significant contribution to understanding teachers' job satisfaction, and offer a methodological approach that can be replicated in similar contexts. Models such as LightGBM and Random Forest can assist decision-makers in the educational sector by identifying teachers with lower satisfaction levels and taking proactive measures to improve their working conditions and overall well-being. This, in turn, would help to enhance educational quality and reinforce the sustainability of the education system. To achieve this, educational administrators should implement strategies such as creating an

optimal work environment, offering continuous training programs, and providing spaces dedicated to reflecting on pedagogical practices. Improving teachers' job satisfaction will not only benefit the educators themselves but will also positively impact students' academic performance and the cohesion of the educational community.

Author Contributions: Conceptualization, L.A.H.-A., N.J.U.-G., R.N.A.-N. and E.E.C.-V.; methodology, L.A.H.-A., E.E.C.-V., F.H.V.-S., T.P.A.-R. and D.D.C.-A.; software, L.A.H.-A., E.E.C.-V., F.H.V.-S., T.P.A.-R., D.D.C.-A. and N.J.U.-G.; validation, L.A.H.-A., E.E.C.-V., R.N.A.-N., R.R.-R., N.K.O.-A. and M.Q.-L.; formal analysis, L.A.H.-A., E.E.C.-V., F.H.V.-S., T.P.A.-R. and N.J.U.-G.; investigation, L.A.H.-A., N.J.U.-G., R.N.A.-N., N.K.O.-A. and E.E.C.-V.; data curation, L.A.H.-A., E.E.C.-V., F.H.V.-S., T.P.A.-R., D.D.C.-A. and N.J.U.-G.; writing—original draft preparation, L.A.H.-A., N.J.U.-G., R.N.A.-N. and E.E.C.-V.; writing—review and editing, L.A.H.-A., M.Q.-L., R.R.-R., N.J.U.-G. and R.N.A.-N.; visualization, L.A.H.-A. and E.E.C.-V.; supervision, L.A.H.-A.; project administration, L.A.H.-A., M.Q.-L. and R.R.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universidad Nacional Amazónica de Madre de Dios, grant number (2023-1GI-16) and The APC was funded by Universidad Nacional Amazónica de Madre de Dios.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are available upon request by contacting the corresponding author, Luis Alberto Holgado-Apaza, via email at lholgado@unamad.edu.pe. Interested parties are encouraged to reach out to the author to obtain access to the data.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Data dictionary of the features selected by the ensemble of methods.

Feature	Description
P1_24_B	Satisfaction with your health
P1_24_C	Satisfaction with the living conditions you can provide for your children/family
P1_24_E	Satisfaction with your job at the educational institution
P1_24_F	Satisfaction with the conditions for carrying out their teaching duties
P1_22_A	Degree of trust with the Ministry of Education
P1_2	Age
P1_22_D	Degree of trust with the Local Management Unit (UGEL)
P1_26_C	Reflection on the results of their pedagogical practice
P1_26_E	Participation in continuing education programs
P1_9_D_LV_HORA	Hours dedicated to household chores and childcare/parental care from Monday to Friday
P1_9_A_SD_HORA	Hours spent on class preparation and administrative tasks on Saturdays and Sundays
P1_6	Number of students under your care
P1_27_E	Difficulty in planning activities under the competency-based approach of the National Basic Education Curriculum
P1_22_C	Level of trust in the Regional Education Directorate or Management
P1_26_B	Difficulty in systematizing pedagogical practice
P1_9_D_SD_HORA	Hours spent on housework and child/parent care on a Saturday and Sunday
P1_9_E_SD_HORA	Hours devoted to leisure or sports (excluding sleep) on a Saturday and Sunday
P1_9_A_LV_HORA	Hours spent preparing classes and administrative tasks from Monday to Friday

References

1. Dagli, A.; Baysal, N. Investigating Teachers' Life Satisfaction. *Univers. J. Educ. Res.* **2017**, *5*, 1250–1256. [[CrossRef](#)]
2. Diener, E.; Emmons, R.A.; Larsen, R.J.; Griffin, S. The Satisfaction with Life Scale. *J. Personal. Assess.* **2010**, *49*, 71–75. [[CrossRef](#)] [[PubMed](#)]
3. Diener, E.; Sapyta, J.J.; Suh, E. Subjective Well-Being Is Essential to Well-Being. *Psychol. Inq.* **2009**, *9*, 33–37. [[CrossRef](#)]
4. Lind, N. Better Life Index. In *Encyclopedia of Quality of Life and Well-Being Research*; Springer: Dordrecht, The Netherlands, 2014; pp. 381–382. [[CrossRef](#)]

5. Helliwell, J.F.; Huang, H.; Shiple, H.; Wang, S. Happiness of the Younger, the Older, and Those in Between. *World Happiness Rep.* **2024**, *2024*, 9–60.
6. UNDP. Human Development Index (HDI) by Country 2024. Available online: <https://worldpopulationreview.com/country-rankings/hdi-by-country> (accessed on 11 August 2024).
7. Malvaso, A.; Kang, W. The Relationship between Areas of Life Satisfaction, Personality, and Overall Life Satisfaction: An Integrated Account. *Front. Psychol.* **2022**, *13*, 894610. [[CrossRef](#)]
8. Angelini, V.; Cavapozzi, D.; Corazzini, L.; Paccagnella, O. Age, Health and Life Satisfaction Among Older Europeans. *Soc. Indic. Res.* **2012**, *105*, 293–308. [[CrossRef](#)]
9. Hong, Y.Z.; Su, Y.J.; Chang, H.H. Analyzing the Relationship between Income and Life Satisfaction of Forest Farm Households—a Behavioral Economics Approach. *For. Policy Econ.* **2023**, *148*, 102916. [[CrossRef](#)]
10. Joshanloo, M.; Jovanović, V. The Relationship between Gender and Life Satisfaction: Analysis across Demographic Groups and Global Regions. *Arch. Women's Ment. Health* **2020**, *23*, 331–338. [[CrossRef](#)]
11. Rogowska, A.M.; Meres, H. The Mediating Role of Job Satisfaction in the Relationship between Emotional Intelligence and Life Satisfaction among Teachers during the COVID-19 Pandemic. *Eur. J. Investig. Health Psychol. Educ.* **2022**, *12*, 666–676. [[CrossRef](#)]
12. Kida, H.; Niimura, H.; Eguchi, Y.; Suzuki, K.; Shikimoto, R.; Bun, S.; Takayama, M.; Mimura, M. Relationship Between Life Satisfaction and Psychological Characteristics Among Community-Dwelling Oldest-Old: Focusing on Erikson's Developmental Stages and the Big Five Personality Traits. *Am. J. Geriatr. Psychiatry* **2024**, *32*, 724–735. [[CrossRef](#)]
13. Kuykendall, L.; Tay, L.; Ng, V. Leisure Engagement and Subjective Well-Being: A Meta-Analysis. *Psychol. Bull.* **2015**, *141*, 364–403. [[CrossRef](#)] [[PubMed](#)]
14. Znidaršič, J.; Marič, M. Relationships between Work-Family Balance, Job Satisfaction, Life Satisfaction and Work Engagement among Higher Education Lecturers. *Organizacija* **2021**, *54*, 227–237. [[CrossRef](#)]
15. Liu, Y.S.; Lu, C.W.; Chung, H.T.; Wang, J.K.; Su, W.J.; Chen, C.W. Health-Promoting Lifestyle and Life Satisfaction in Full-Time Employed Adults with Congenital Heart Disease: Grit as a Mediator. *Eur. J. Cardiovasc. Nurs.* **2024**, *23*, 348–357. [[CrossRef](#)]
16. Kim, E.-J.; Kang, H.-W.; Sala, A.; Kim, E.-J.; Kang, H.-W.; Park, S.-M. Leisure and Happiness of the Elderly: A Machine Learning Approach. *Sustainability* **2024**, *16*, 2730. [[CrossRef](#)]
17. Phulkerd, S.; Thapsuwan, S.; Chamratthirong, A.; Gray, R.S. Influence of Healthy Lifestyle Behaviors on Life Satisfaction in the Aging Population of Thailand: A National Population-Based Survey. *BMC Public Health* **2021**, *21*, 43. [[CrossRef](#)]
18. Zagkas, D.G.; Chrousos, G.P.; Bacopoulou, F.; Kanaka-Gantenbein, C.; Vlachakis, D.; Tzelepi, I.; Darviri, C. Stress and Well-Being of Greek Primary School Educators: A Cross-Sectional Study. *Int. J. Environ. Res. Public Health* **2023**, *20*, 5390. [[CrossRef](#)]
19. Pagán-Castaño, E.; Sánchez-García, J.; Garrigos-Simon, F.J.; Guijarro-García, M. The Influence of Management on Teacher Well-Being and the Development of Sustainable Schools. *Sustainability* **2021**, *13*, 2909. [[CrossRef](#)]
20. Ao, N.; Zhang, S.; Tian, G.; Zhu, X.; Kang, X. Exploring Teacher Wellbeing in Educational Reforms: A Chinese Perspective. *Front. Psychol.* **2023**, *14*, 1265536. [[CrossRef](#)] [[PubMed](#)]
21. Natha, P.; RajaRajeswari, P. Advancing Skin Cancer Prediction Using Ensemble Models. *Computers* **2024**, *13*, 157. [[CrossRef](#)]
22. Conte, L.; De Nunzio, G.; Giombi, F.; Lupo, R.; Arigliani, C.; Leone, F.; Salamanca, F.; Petrelli, C.; Angelelli, P.; De Benedetto, L.; et al. Machine Learning Models to Enhance the Berlin Questionnaire Detection of Obstructive Sleep Apnea in At-Risk Patients. *Appl. Sci.* **2024**, *14*, 5959. [[CrossRef](#)]
23. Ghassemi, M.; Naumann, T.; Schulam, P.; Beam, A.L.; Chen, I.Y.; Ranganath, R. A Review of Challenges and Opportunities in Machine Learning for Health. *AMIA Summits Transl. Sci. Proc.* **2020**, *2020*, 191. [[PubMed](#)]
24. Stephen, O.; Sain, M.; Maduh, U.J.; Jeong, D.U. An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare. *J. Healthc. Eng. J.* **2019**, *2019*, 4180949. [[CrossRef](#)]
25. Spencer, R.; Thabtah, F.; Abdelhamid, N.; Thompson, M. Exploring Feature Selection and Classification Methods for Predicting Heart Disease. *Digit. Heal.* **2020**, *6*, 2055207620914777. [[CrossRef](#)] [[PubMed](#)]
26. Hamdia, K.M.; Zhuang, X.; Rabczuk, T. An Efficient Optimization Approach for Designing Machine Learning Models Based on Genetic Algorithm. *Neural Comput. Appl.* **2021**, *33*, 1923–1933. [[CrossRef](#)]
27. Bhosekar, A.; Ierapetritou, M. Modular Design Optimization Using Machine Learning-Based Flexibility Analysis. *J. Process Control* **2020**, *90*, 18–34. [[CrossRef](#)]
28. Yogesh, I.; Suresh Kumar, K.R.; Candrashekar, N.; Reddy, D.; Sampath, H. Predicting Job Satisfaction and Employee Turnover Using Machine Learning. *J. Comput. Theor. Nanosci.* **2020**, *17*, 4092–4097. [[CrossRef](#)]
29. Celbiş, M.G.; Wong, P.H.; Kourtit, K.; Nijkamp, P. Job Satisfaction and the 'Great Resignation': An Exploratory Machine Learning Analysis. *Soc. Indic. Res.* **2023**, *170*, 1097–1118. [[CrossRef](#)]
30. Gupta, A.; Chadha, A.; Tiwari, V.; Varma, A.; Pereira, V. Sustainable Training Practices: Predicting Job Satisfaction and Employee Behavior Using Machine Learning Techniques. *Asian Bus. Manag.* **2023**, *22*, 1913–1936. [[CrossRef](#)]
31. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [[CrossRef](#)]
32. Sharma, A.; Jain, A.; Gupta, P.; Chowdary, V. Machine Learning Applications for Precision Agriculture: A Comprehensive Review. *IEEE Access* **2021**, *9*, 4843–4873. [[CrossRef](#)]
33. Benos, L.; Tagarakis, A.C.; Dolias, G.; Berruto, R.; Kateris, D.; Bochtis, D. Machine Learning in Agriculture: A Comprehensive Updated Review. *Sensors* **2021**, *21*, 3758. [[CrossRef](#)] [[PubMed](#)]

34. Pallathadka, H.; Mustafa, M.; Sanchez, D.T.; Sekhar Sajja, G.; Gour, S.; Naved, M. Impact of Machine Learning on Management, Healthcare and Agriculture. *Mater. Today Proc.* **2023**, *80*, 2803–2806. [CrossRef]
35. McQueen, R.J.; Garner, S.R.; Nevill-Manning, C.G.; Witten, I.H. Applying Machine Learning to Agricultural Data. *Comput. Electron. Agric.* **1995**, *12*, 275–293. [CrossRef]
36. Leo, M.; Sharma, S.; Maddulety, K. Machine Learning in Banking Risk Management: A Literature Review. *Risks* **2019**, *7*, 29. [CrossRef]
37. Mashrur, A.; Luo, W.; Zaidi, N.A.; Robles-Kelly, A. Machine Learning for Financial Risk Management: A Survey. *IEEE Access* **2020**, *8*, 203203–203223. [CrossRef]
38. Aziz, S.; Dowling, M.M. AI and Machine Learning for Risk Management. *SSRN Electron. J.* **2018**, 33–50. [CrossRef]
39. Mandapuram, M.; Mandapuram, M.; Gutlapalli, S.S.; Reddy, M.; Bodepudi, A. Application of Artificial Intelligence (AI) Technologies to Accelerate Market Segmentation. *Glob. Discl. Econ. Bus.* **2020**, *9*, 141–150. [CrossRef]
40. Ngai, E.W.T.; Wu, Y. Machine Learning in Marketing: A Literature Review, Conceptual Framework, and Research Agenda. *J. Bus. Res.* **2022**, *145*, 35–48. [CrossRef]
41. Yoganarasimhan, H. Search Personalization Using Machine Learning. *Manag. Sci.* **2019**, *66*, 1045–1070. [CrossRef]
42. Greene, T.; Shmueli, G. How Personal Is Machine Learning Personalization? *arXiv* **2019**, arXiv:1912.07938.
43. Lovera, F.A.; Cardinale, Y. Sentiment Analysis in Twitter: A Comparative Study. *Rev. Cient. Sist. E Informática* **2023**, *3*, e418. [CrossRef]
44. Sentieiro, D.H. Machine Learning for Autonomous Vehicle Route Planning and Optimization. *J. AI-Assist. Sci. Discov.* **2022**, *2*, 1–20.
45. Lazar, D.A.; Bıyık, E.; Sadigh, D.; Pedarsani, R. Learning How to Dynamically Route Autonomous Vehicles on Shared Roads. *Transp. Res. Part C: Emerg. Technol.* **2021**, *130*, 103258. [CrossRef]
46. Lee, S.; Kim, Y.; Kahng, H.; Lee, S.K.; Chung, S.; Cheong, T.; Shin, K.; Park, J.; Kim, S.B. Intelligent Traffic Control for Autonomous Vehicle Systems Based on Machine Learning. *Expert Syst. Appl.* **2020**, *144*, 113074. [CrossRef]
47. Liu, Y.; Fan, S.; Xu, S.; Sajjanhar, A.; Yeom, S.; Wei, Y. Predicting Student Performance Using Clickstream Data and Machine Learning. *Educ. Sci.* **2022**, *13*, 17. [CrossRef]
48. Alghamdi, A.S.; Rahman, A. Data Mining Approach to Predict Success of Secondary School Students: A Saudi Arabian Case Study. *Educ. Sci.* **2023**, *13*, 293. [CrossRef]
49. Bayazit, A.; Apaydin, N.; Gonullu, I. Predicting At-Risk Students in an Online Flipped Anatomy Course Using Learning Analytics. *Educ. Sci.* **2022**, *12*, 581. [CrossRef]
50. Zhang, C.; Ahn, H. E-Learning at-Risk Group Prediction Considering the Semester and Realistic Factors. *Educ. Sci.* **2023**, *13*, 1130. [CrossRef]
51. MINEDU. Ministerio de Educación del Perú | MINEDU. Available online: <http://www.minedu.gob.pe/politicas/docencia/encuesta-nacional-a-docentes-endo.php> (accessed on 8 May 2021).
52. Diener, E.; Diener, M. Cross-Cultural Correlates of Life Satisfaction and Self-Esteem. *J. Personal. Soc. Psychol.* **1995**, *68*, 653–663. [CrossRef]
53. Karataş, Z.; Uzun, K.; Tagay, Ö. Relationships Between the Life Satisfaction, Meaning in Life, Hope and COVID-19 Fear for Turkish Adults During the COVID-19 Outbreak. *Front. Psychol.* **2021**, *12*, 633384. [CrossRef]
54. Szcześniak, M.; Tułeczka, M. Family Functioning and Life Satisfaction: The Mediator Role of Emotional Intelligence. *Psychol. Res. Behav. Manag.* **2020**, *13*, 223–232. [CrossRef]
55. Wang, Y.; Zhang, D. Economic Income and Life Satisfaction of Rural Chinese Older Adults: The Effects of Physical Health and Ostracism. *Research Square* **2022**. [CrossRef]
56. Judge, T.A.; Piccolo, R.F.; Podsakoff, N.P.; Shaw, J.C.; Rich, B.L. The Relationship between Pay and Job Satisfaction: A Meta-Analysis of the Literature. *J. Vocat. Behav.* **2010**, *77*, 157–167. [CrossRef]
57. Haar, J.M.; Russo, M.; Suñe, A.; Ollier-Malaterre, A. Outcomes of Work–Life Balance on Job Satisfaction, Life Satisfaction and Mental Health: A Study across Seven Cultures. *J. Vocat. Behav.* **2014**, *85*, 361–373. [CrossRef]
58. Noda, H. Work–Life Balance and Life Satisfaction in OECD Countries: A Cross-Sectional Analysis. *J. Happiness Stud.* **2020**, *21*, 1325–1348. [CrossRef]
59. Author, C.; Hee Park, K. The Relationships between Well-Being Lifestyle, Well-Being Attitude, Life Satisfaction, and Demographic Characteristics. *J. Korean Home Econ. Assoc.* **2011**, *49*, 39–49. [CrossRef]
60. Luque-Reca, O.; García-Martínez, I.; Pulido-Martos, M.; Lorenzo Burguera, J.; Augusto-Landa, J.M. Teachers’ Life Satisfaction: A Structural Equation Model Analyzing the Role of Trait Emotion Regulation, Intrinsic Job Satisfaction and Affect. *Teach. Teach. Educ.* **2022**, *113*, 103668. [CrossRef]
61. Lent, R.W.; Nota, L.; Soresi, S.; Ginevra, M.C.; Duffy, R.D.; Brown, S.D. Predicting the Job and Life Satisfaction of Italian Teachers: Test of a Social Cognitive Model. *J. Vocat. Behav.* **2011**, *79*, 91–97. [CrossRef]
62. Cayupe, J.C.; Bernedo-Moreira, D.H.; Morales-García, W.C.; Alcaraz, F.L.; Peña, K.B.C.; Saintila, J.; Flores-Paredes, A. Self-Efficacy, Organizational Commitment, Workload as Predictors of Life Satisfaction in Elementary School Teachers: The Mediating Role of Job Satisfaction. *Front. Psychol.* **2023**, *14*, 1066321. [CrossRef]
63. Marcionetti, J.; Castelli, L. The Job and Life Satisfaction of Teachers: A Social Cognitive Model Integrating Teachers’ Burnout, Self-Efficacy, Dispositional Optimism, and Social Support. *Int. J. Educ. Vocat. Guid.* **2023**, *23*, 441–463. [CrossRef]

64. Bano, S.; Malik, S.; Sadia, M. Effect of Occupational Stress on Life Satisfaction among Private and Public School Teachers. *JISR Manag. Soc. Sci. Econ.* **2014**, *12*, 61–72. [[CrossRef](#)]
65. Quinteros-Durand, R.; Almanza-Cabe, R.B.; Morales-García, W.C.; Mamani-Benito, O.; Sairitupa-Sanchez, L.Z.; Puño-Quispe, L.; Saintila, J.; Saavedra-Sandoval, R.; Paredes, A.F.; Ramírez-Coronel, A.A. Influence of Servant Leadership on the Life Satisfaction of Basic Education Teachers: The Mediating Role of Satisfaction with Job Resources. *Front. Psychol.* **2023**, *14*, 1167074. [[CrossRef](#)]
66. Sanchez-Martinez, S.; Camara, O.; Piella, G.; Cikes, M.; González-Ballester, M.Á.; Miron, M.; Vellido, A.; Gómez, E.; Fraser, A.G.; Bijmens, B. Machine Learning for Clinical Decision-Making: Challenges and Opportunities in Cardiovascular Imaging. *Front. Cardiovasc. Med.* **2021**, *8*, 765693. [[CrossRef](#)] [[PubMed](#)]
67. Byeon, H. Application of Artificial Neural Network Analysis and Decision Tree Analysis to Develop a Model for Predicting Life Satisfaction of the Elderly in South Korea. *Int. J. Eng. Technol.* **2018**, *7*, 161–166. [[CrossRef](#)]
68. Zhang, J.; Li, L. *A Study on Life Satisfaction Prediction of the Elderly Based on SVM*; Association for Computing Machinery: New York, NY, USA, 2023; pp. 16–21. [[CrossRef](#)]
69. Pan, Z.; Cutumisu, M. Using Machine Learning to Predict UK and Japanese Secondary Students' Life Satisfaction in PISA 2018. *Br. J. Educ. Psychol.* **2024**, *94*, 474–498. [[CrossRef](#)]
70. Khan, A.E.; Hasan, M.J.; Anjum, H.; Mohammed, N.; Momen, S. Predicting Life Satisfaction Using Machine Learning and Explainable AI. *Heliyon* **2024**, *10*, e31158. [[CrossRef](#)]
71. Jaiswal, R.; Gupta, S. Money Talks, Happiness Walks: Dissecting the Secrets of Global Bliss with Machine Learning. *J. Chin. Econ. Bus. Stud.* **2024**, *22*, 111–158. [[CrossRef](#)]
72. Morrone, A.; Piscitelli, A.; D'Ambrosio, A. How Disadvantages Shape Life Satisfaction: An Alternative Methodological Approach. *Soc. Indic. Res.* **2019**, *141*, 477–502. [[CrossRef](#)]
73. Lee, S. Exploring Factors Influencing Life Satisfaction of Youth Using Random Forests. *J. Ind. Converg.* **2023**, *21*, 9–17. [[CrossRef](#)]
74. Shen, X.; Yin, F.; Jiao, C. Predictive Models of Life Satisfaction in Older People: A Machine Learning Approach. *Int. J. Environ. Res. Public Health* **2023**, *20*, 2445. [[CrossRef](#)]
75. Jang, J.H.; Masatsuku, N. A Study of Factors Influencing Happiness in Korea: Topic Modelling and Neural Network Analysis [Estudio de Los Factores Que Influyen En La Felicidad En Corea: Modelización de Temas y Análisis de Redes Neuronales]. *Data Metadata* **2024**, *3*, 238. [[CrossRef](#)]
76. Pudjihartono, N.; Fadason, T.; Kempa-Liehr, A.W.; O'Sullivan, J.M. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Front. Bioinforma.* **2022**, *2*, 927312. [[CrossRef](#)] [[PubMed](#)]
77. Yang, P.; Zhou, B.B.; Yang, J.Y.H.; Zomaya, A.Y. Stability of Feature Selection Algorithms and Ensemble Feature Selection Methods in Bioinformatics. In *Biological Knowledge Discovery Handbook: Preprocessing, Mining and Postprocessing of Biological Data*; John Wiley & Sons, Inc.: New York, NY, USA, 2014; pp. 333–352. [[CrossRef](#)]
78. Abeel, T.; Helleputte, T.; Van de Peer, Y.; Dupont, P.; Saeys, Y. Robust Biomarker Identification for Cancer Diagnosis with Ensemble Feature Selection Methods. *Bioinformatics* **2010**, *26*, 392–398. [[CrossRef](#)] [[PubMed](#)]
79. Wang, J.; Xu, J.; Zhao, C.; Peng, Y.; Wang, H. An Ensemble Feature Selection Method for High-Dimensional Data Based on Sort Aggregation. *Syst. Sci. Control. Eng.* **2019**, *7*, 32–39. [[CrossRef](#)]
80. Tsai, C.F.; Sung, Y.T. Ensemble Feature Selection in High Dimension, Low Sample Size Datasets: Parallel and Serial Combination Approaches. *Knowl.-Based Syst.* **2020**, *203*, 106097. [[CrossRef](#)]
81. Hoque, N.; Singh, M.; Bhattacharyya, D.K. EFS-MI: An Ensemble Feature Selection Method for Classification. *Complex Intell. Syst.* **2017**, *4*, 105–118. [[CrossRef](#)]
82. Seijo-Pardo, B.; Porto-Díaz, I.; Bolón-Canedo, V.; Alonso-Betanzos, A. Ensemble Feature Selection: Homogeneous and Heterogeneous Approaches. *Knowl.-Based Syst.* **2017**, *118*, 124–139. [[CrossRef](#)]
83. Ben Brahim, A.; Limam, M. Ensemble Feature Selection for High Dimensional Data: A New Method and a Comparative Study. *Adv. Data Anal. Classif.* **2018**, *12*, 937–952. [[CrossRef](#)]
84. Neumann, U.; Genze, N.; Heider, D. EFS: An Ensemble Feature Selection Tool Implemented as R-Package and Web-Application. *BioData Min.* **2017**, *10*, 21. [[CrossRef](#)]
85. Werner de Vargas, V.; Schneider Aranda, J.A.; dos Santos Costa, R.; da Silva Pereira, P.R.; Victória Barbosa, J.L. Imbalanced Data Preprocessing Techniques for Machine Learning: A Systematic Mapping Study. *Knowl. Inf. Syst.* **2023**, *65*, 31–57. [[CrossRef](#)]
86. Gardner, W.; Winkler, D.A.; Alexander, D.L.J.; Ballabio, D.; Muir, B.W.; Pigram, P.J. Effect of Data Preprocessing and Machine Learning Hyperparameters on Mass Spectrometry Imaging Models. *J. Vac. Sci. Technol. A* **2023**, *41*, 63204. [[CrossRef](#)]
87. Frye, M.; Mohren, J.; Schmitt, R.H. Benchmarking of Data Preprocessing Methods for Machine Learning-Applications in Production. *Procedia CIRP* **2021**, *104*, 50–55. [[CrossRef](#)]
88. Maharana, K.; Mondal, S.; Nemade, B. A Review: Data Pre-Processing and Data Augmentation Techniques. *Glob. Transit. Proc.* **2022**, *3*, 91–99. [[CrossRef](#)]
89. Dina Diatta, I.; Berchtold, A. Impact of Missing Information on Day-to-Day Research Based on Secondary Data. *Int. J. Soc. Res. Methodol.* **2023**, *26*, 759–772. [[CrossRef](#)]
90. Austin, P.C.; White, I.R.; Lee, D.S.; van Buuren, S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. *Can. J. Cardiol.* **2021**, *37*, 1322–1331. [[CrossRef](#)]
91. Emmanuel, T.; Maupong, T.; Mpoeleng, D.; Semong, T.; Mphago, B.; Tabona, O. A Survey on Missing Data in Machine Learning. *J. Big Data* **2021**, *8*, 140. [[CrossRef](#)] [[PubMed](#)]

92. Memon, S.M.; Wamala, R.; Kabano, I.H. A Comparison of Imputation Methods for Categorical Data. *Inform. Med. Unlocked* **2023**, *42*, 101382. [[CrossRef](#)]
93. Kosaraju, N.; Sankepally, S.R.; Mallikharjuna Rao, K. Categorical Data: Need, Encoding, Selection of Encoding Method and Its Emergence in Machine Learning Models—A Practical Review Study on Heart Disease Prediction Dataset Using Pearson Correlation. *Lect. Notes Networks Syst.* **2023**, *1*, 369–382. [[CrossRef](#)]
94. Mallikharjuna Rao, K.; Saikrishna, G.; Supriya, K. Data Preprocessing Techniques: Emergence and Selection towards Machine Learning Models—A Practical Review Using HPA Dataset. *Multimed. Tools Appl.* **2023**, *82*, 37177–37196. [[CrossRef](#)]
95. Vowels, L.M.; Vowels, M.J.; Mark, K.P. Identifying the Strongest Self-Report Predictors of Sexual Satisfaction Using Machine Learning. *J. Soc. Pers. Relat.* **2022**, *39*, 1191–1212. [[CrossRef](#)]
96. Zhang, H.; Zheng, G.; Xu, J.; Yao, X. Research on the Construction and Realization of Data Pipeline in Machine Learning Regression Prediction. *Math. Probl. Eng.* **2022**, *2022*, 7924335. [[CrossRef](#)]
97. Md, A.Q.; Kulkarni, S.; Joshua, C.J.; Vaichole, T.; Mohan, S.; Iwendi, C. Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease. *Biomedicines* **2023**, *11*, 581. [[CrossRef](#)] [[PubMed](#)]
98. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
99. Daza Vergaray, A.; Miranda, J.C.H.; Cornelio, J.B.; López Carranza, A.R.; Ponce Sánchez, C.F. Predicting the Depression in University Students Using Stacking Ensemble Techniques over Oversampling Method. *Inform. Med. Unlocked* **2023**, *41*, 101295. [[CrossRef](#)]
100. Lemaître, G.; Nogueira, F.; Aridas, C.K. Imbalanced-Learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *J. Mach. Learn. Res.* **2017**, *18*, 1–5.
101. Wang, L.; Han, M.; Li, X.; Zhang, N.; Cheng, H. Review of Classification Methods on Unbalanced Data Sets. *IEEE Access* **2021**, *9*, 64606–64628. [[CrossRef](#)]
102. Vilorio, A.; Lezama, O.B.P.; Mercado-Caruzo, N. Unbalanced Data Processing Using Oversampling: Machine Learning. *Procedia Comput. Sci.* **2020**, *175*, 108–113. [[CrossRef](#)]
103. Xu, Y.; Park, Y.; Park, J.D.; Sun, B. Predicting Nurse Turnover for Highly Imbalanced Data Using the Synthetic Minority Over-Sampling Technique and Machine Learning Algorithms. *Healthcare* **2023**, *11*, 3173. [[CrossRef](#)]
104. Kalimuthan, C.; Arokia Renjit, J. Review on Intrusion Detection Using Feature Selection with Machine Learning Techniques. *Mater. Today Proc.* **2020**, *33*, 3794–3802. [[CrossRef](#)]
105. Abubakar, S.M.; Sufyanu, Z.; Abubakar, M.M. A survey of feature selection methods for software defect prediction models. *FUDMA J. Sci.* **2020**, *4*, 62–68.
106. Chandrashekar, G.; Sahin, F. A Survey on Feature Selection Methods. *Comput. Electr. Eng.* **2014**, *40*, 16–28. [[CrossRef](#)]
107. Miao, J.; Niu, L. A Survey on Feature Selection. *Procedia Comput. Sci.* **2016**, *91*, 919–926. [[CrossRef](#)]
108. Jia, W.; Sun, M.; Lian, J.; Hou, S. Feature Dimensionality Reduction: A Review. *Complex Intell. Syst.* **2022**, *8*, 2663–2693. [[CrossRef](#)]
109. Altae, A.A.; Rad, A.E.; Tati, R. Comparative Study on Effective Feature Selection Methods. *Int. J. Innov. Eng. Manag. Res. Forthcoming*. **2023**.
110. Tang, J.; Alelyani, S.; Liu, H. Feature Selection for Classification: A Review. *Data Classif. Algorithms Appl.* **2014**, 37–64. [[CrossRef](#)]
111. Bommert, A.; Sun, X.; Bischl, B.; Rahnenführer, J.; Lang, M. Benchmark for Filter Methods for Feature Selection in High-Dimensional Classification Data. *Comput. Stat. Data Anal.* **2020**, *143*, 106839. [[CrossRef](#)]
112. Nguyen, H.B.; Xue, B.; Andreae, P. Mutual Information Estimation for Filter Based Feature Selection Using Particle Swarm Optimization. In *Applications of Evolutionary Computation, Proceedings of the 19th European Conference, Ponto, Portugal, 30 March–1 April 2016*; pp. 719–736. [[CrossRef](#)]
113. Vergara, J.R.; Estévez, P.A. A Review of Feature Selection Methods Based on Mutual Information. *Neural Comput. Appl.* **2014**, *24*, 175–186. [[CrossRef](#)]
114. Dissanayake, K.; Johar, M.G.M. Comparative Study on Heart Disease Prediction Using Feature Selection Techniques on Classification Algorithms. *Appl. Comput. Intell. Soft Comput.* **2021**, *2021*, 5581806. [[CrossRef](#)]
115. Tripathy, G.; Sharaff, A. AEGA: Enhanced Feature Selection Based on ANOVA and Extended Genetic Algorithm for Online Customer Review Analysis. *J. Supercomput.* **2023**, *79*, 13180–13209. [[CrossRef](#)]
116. Raufi, B.; Longo, L. Comparing ANOVA and PowerShap Feature Selection Methods via Shapley Additive Explanations of Models of Mental Workload Built with the Theta and Alpha EEG Band Ratios. *BioMedInformatics* **2024**, *4*, 853–876. [[CrossRef](#)]
117. Laborda, J.; Ryoo, S. Feature Selection in a Credit Scoring Model. *Mathematics* **2021**, *9*, 746. [[CrossRef](#)]
118. Jiang, J.; Zhang, X.; Yuan, Z. Feature Selection for Classification with Spearman’s Rank Correlation Coefficient-Based Self-Information in Divergence-Based Fuzzy Rough Sets. *Expert Syst. Appl.* **2024**, *249*, 123633. [[CrossRef](#)]
119. Tang, M.; Zhao, Q.; Wu, H.; Wang, Z. Cost-Sensitive LightGBM-Based Online Fault Detection Method for Wind Turbine Gearboxes. *Front. Energy Res.* **2021**, *9*, 701574. [[CrossRef](#)]
120. Liu, H.; Zhou, M.; Liu, Q. An Embedded Feature Selection Method for Imbalanced Data Classification. *IEEE/CAA J. Autom. Sin.* **2019**, *6*, 703–715. [[CrossRef](#)]
121. Elgeldawi, E.; Sayed, A.; Galal, A.R.; Zaki, A.M. Hyperparameter Tuning for Machine Learning Algorithms Used for Arabic Sentiment Analysis. *Informatics* **2021**, *8*, 79. [[CrossRef](#)]
122. Papernot, N.; Steinke, T. Hyperparameter Tuning with Renyi Differential Privacy. *arXiv* **2021**, arXiv:2110.03620.

123. Bacanin, N.; Stoean, C.; Zivkovic, M.; Rakic, M.; Strulak-Wójcikiewicz, R.; Stoean, R. On the Benefits of Using Metaheuristics in the Hyperparameter Tuning of Deep Learning Models for Energy Load Forecasting. *Energies* **2023**, *16*, 1434. [[CrossRef](#)]
124. Ali, Y.A.; Awwad, E.M.; Al-Razgan, M.; Maarouf, A. Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity. *Processes* **2023**, *11*, 349. [[CrossRef](#)]
125. Rajendran, S.; Chamundeswari, S.; Sinha, A.A. Predicting the Academic Performance of Middle- and High-School Students Using Machine Learning Algorithms. *Soc. Sci. Humanit. Open* **2022**, *6*, 100357. [[CrossRef](#)]
126. Passos, D.; Mishra, P. A Tutorial on Automatic Hyperparameter Tuning of Deep Spectral Modelling for Regression and Classification Tasks. *Chemom. Intell. Lab. Syst.* **2022**, *223*, 104520. [[CrossRef](#)]
127. Bergstra, J.; Ca, J.B.; Ca, Y.B. Random Search for Hyper-Parameter Optimization Yoshua Bengio. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
128. Heydarian, M.; Doyle, T.E.; Samavi, R. MLCM: Multi-Label Confusion Matrix. *IEEE Access* **2022**, *10*, 19083–19095. [[CrossRef](#)]
129. Prusty, S.; Patnaik, S.; Dash, S.K. SKCV: Stratified K-Fold Cross-Validation on ML Classifiers for Predicting Cervical Cancer. *Front. Nanotechnol.* **2022**, *4*, 972421. [[CrossRef](#)]
130. Li, X.; Lin, X.; Zhang, F.; Tian, Y. Playing Roles in Work and Family: Effects of Work/Family Conflicts on Job and Life Satisfaction Among Junior High School Teachers. *Front. Psychol.* **2021**, *12*, 772025. [[CrossRef](#)]
131. Judge, T.A.; Bono, J.E. Relationship of Core Self-Evaluations Traits - Self-Esteem, Generalized Self-Efficacy, Locus of Control, and Emotional Stability—With Job Satisfaction and Job Performance: A Meta-Analysis. *J. Appl. Psychol.* **2001**, *86*, 80–92. [[CrossRef](#)] [[PubMed](#)]
132. Holgado-Apaza, L.A.; Carpio-Vargas, E.E.; Calderon-Vilca, H.D.; Maquera-Ramirez, J.; Ulloa-Gallardo, N.J.; Acosta-Navarrete, M.S.; Barrón-Adame, J.M.; Quispe-Layme, M.; Hidalgo-Pozzi, R.; Valles-Coral, M. Modeling Job Satisfaction of Peruvian Basic Education Teachers Using Machine Learning Techniques. *Appl. Sci.* **2023**, *13*, 3945. [[CrossRef](#)]
133. Cole, C.; Hinchcliff, E.; Carling, R. Reflection as Teachers: Our Critical Developments. *Front. Educ.* **2022**, *7*, 1037280. [[CrossRef](#)]
134. Shandomo, H.M. The Role of Critical Reflection in Teacher Education. *Sch.-Univ. Partnersh.* **2010**, *4*, 101–113.
135. Shiri, R.; El-Metwally, A.; Sallinen, M.; Pöyry, M.; Härmä, M.; Toppinen-Tanner, S. The Role of Continuing Professional Training or Development in Maintaining Current Employment: A Systematic Review. *Healthcare* **2023**, *11*, 2900. [[CrossRef](#)]
136. Law, S.F.; Le, A.T. A Systematic Review of Empirical Studies on Trust between Universities and Society. *J. High. Educ. Policy Manag.* **2023**, *45*, 393–408. [[CrossRef](#)]
137. OECD. *OECD Survey on Drivers of Trust in Public Institutions-2024 Results Building Trust In A Complex Policy Environment*; OECD: Paris, France, 2024. [[CrossRef](#)]
138. Helliwell, J.F.; Huang, H. New measures of the costs of unemployment: Evidence from the subjective well-being of 3.3 million americans. *Econ. Inq.* **2014**, *52*, 1485–1502. [[CrossRef](#)]
139. Helliwell, J.; Layard, R.; Sachs, J.; Neve De, J.-E.; Akinin, L. Happiness and Age: Summary | The World Happiness Report. Available online: <https://worldhappiness.report/ed/2024/happiness-and-age-summary/> (accessed on 15 August 2024).
140. Cho, H.; Pyun, D.Y.; Wang, C.K.J. Teachers' Work-Life Balance: The Effect of Work-Leisure Conflict on Work-Related Outcomes. *Asia Pac. J. Educ.* **2023**, 1–16. [[CrossRef](#)]
141. Ertürk, R. The Effect of Teachers' Quality of Work Life on Job Satisfaction and Turnover Intentions. *Int. J. Contemp. Educ. Res.* **2022**, *9*, 191–203. [[CrossRef](#)]
142. Lee, K.O.; Lee, K.S. Effects of Emotional Labor, Anger, and Work Engagement on Work-Life Balance of Mental Health Specialists Working in Mental Health Welfare Centers. *Int. J. Environ. Res. Public Heal.* **2023**, *20*, 2353. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.